DOCUMENT RESUME

ED 334 969 IR 015 132

AUTHOR	Simonson, Michael R., Ed.; Hargrave, Connie, Ed.
TITLE	Proceedings of Selected Research Paper Presentations at the Convention of the Association for Educational Communications and Technology and Sponsored by the Research and Theory Division (Orlando, Florida, February 13-17, 1991).
INSTITUTION	Iowa State Univ. of Science and Technology, Ames. Coll. of Education.
PUB DATE	91
NOTE	1,025p.; For the individual papers, see IR 015 133-153, IR 015 158-193, and ED 323 937. For the proceedings of the 1990 conference, see ED 323 912.
PUB TYPE	Collected Works - Conference Proceedings (021) -- Reports - Research/Technical (143)
EDRS PRICE	MF07/PC41 Plus Postage.
DESCRIPTORS	Cognitive Processes; *Computer Assisted Instruction; Cooperative Learning; Distance Education; Educational Technology; Elementary Secondary Education; Epistemology; Expert Systems; Higher Education; Hypermedia; Industrial Training; *Instructional Design; Instructional Development; *Instructional Effectiveness; Interactive Video; *Learner Controlled Instruction; Locus of Control; Media Research; Models; *Research Methodology

ABSTRACT

This volume contains 57 papers and 1 symposium, consisting of an introduction and 4 papers. The topics discussed range from the instructional effectiveness of various presentation formats and media to the social foundations of educational computing and school restructuring. Cognitive theory and several design theories are also addressed. Research related to computer assisted instruction includes investigations of cooperative learning, variations in instructional control, interactive video, use of computer animation, hypermedia browsers, software for teachers to use in lesson planning, computer assisted testing, and automated instructional systems development. Formats for microcomputer displays and electronic publishing are also considered, as well as computer simulations and telecommunications-based distance education. Also included in this volume are information on the Association for Educational Communications and Technology (AECT), author and descriptor indexes, and a list of the 12 proceedings already in the ERIC system together with their ED numbers. (BBM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ED334969

13th ANNUAL

PROCEEDINGS

of
SELECTED
RESEARCH PRESENTATIONS

at the
1991 Annual Convention
of the
Association for
Educational Communications
and Technology

Sponsored by the
Research and Theory Division
Orlando, FL 1991

aect

EDITORS:
Michael R. Simonson and Connie Hargrave
Iowa State University
Ames, Iowa

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Michael Simonson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR015132

13th ANNUAL PROCEEDINGS of SELECTED RESEARCH PRESENTATIONS

at the 1991 Convention of the
Association for Educational Communications and Technology -
Sponsored by the Research and Theory Division
in
Orlando, Florida

Edited by:

Michael R. Simonson
Professor of Secondary Education

and

Connie Hargrave
Teaching Assistant

Iowa State University
College of Education
Instructional Resources Center
Lagomarcino Hall
Ames, Iowa 50011

(515) 294-6840

PREFACE

For the thirteenth year, the Research and Theory Division of the Association for Educational Communications and Technology (AECT) is sponsoring the publications of these Proceedings. Papers published in this volume were presented at the national AECT Convention in Orlando, Florida. A limited quantity of this volume were printed and sold. It is also available on microfiche through the Educational Resources Information Clearinghouse (ERIC) system.

REFEREEING PROCESS: All research papers selected for presentation at the AECT convention and included in this Proceedings were subjected to a rigorous blind reviewing process. Proposals were submitted to Steven Ross and Gary Morrison of Memphis State University or Marina McIsaac of Arizona State University. All references to author were removed from proposals before they were submitted to referees for review. Approximately fifty percent of the manuscripts submitted for consideration were selected for presentation at the Convention and for Publication in these Proceedings. The papers contained in this document represent some of the most current thinking in educational communications and technology.

This volume contains two indexes covering this volume. The first is an author index; the second is a descriptor index. The index for volumes 1-6 (1979-84) is included in the 1986 Proceedings, and the index for volumes 7-10 is in the 1988 Proceedings.

M. R. Simonson
Editor

RESEARCH AND THEORY DIVISION OFFICERS

Past President

John Wedman
1110 Parkridge Dr.
Columbia, MO 65203
314-882-3828

President

Marina McIsaac
Educational Media & Computers
Arizona State University
College of Education, Payne 146
Tempe, AZ 85287-0111
602-965-4961

President Elect

Gary Anglin
315 Stoneybrook Dr.
Lexington, KY 40503
606-258-5972

Board Members

Dean Christensen
BLCS1F
Control Data Corporation
P.O. Box 1305
Minneapolis, MN 55440
612-921-7099

Randy Koetting
College of Education
Oklahoma State University
Stillwater, OK 74078

Gary Morrison
Department of Curriculum & Instruction
Education Building 424
Memphis State University
Memphis, Tennessee 38152

Charles Reigeluth
Professor
Indiana University
210 Education Building
Bloomington, IN 47405
812-855-6118

Rhonda Robinson
Associate Professor
Northern Illinois University
LEPS-GA219
De Kalb, IL 60115
815-753-9323

Brent Wilson
Associate Professor
University of Colorado
1200 Larimer, Box 106
Denver, CO 80204
303-556-4363
303-556-4822

Andrew Yeaman
Yeaman & Associates
Apt. 12
222 N. Broad St.
Monmouth, OR 97361
503-838-5649
503-838-8474

Association for Educational Communications and Technology

What is AECT?

AECT is the only national, professional association dedicated to the improvement of instruction through the effective use of media and technology. AECT assists its members in using technology in their jobs and to enhance the learning process.

Who belongs to AECT?

○ Media Specialists
○ Educators
○ Instructional Designers
○ Learning Resource Specialists
○ Curriculum Developers
○ Television producers and directors
○ Communications specialists
○ Education Administrators
○ Others who require expertise in instructional technology

AECT History

AECT began as the Department of Audio-Visual Instruction at the National Education Association in 1923, in the days when visual aids were blackboards and easels. In 1947, as educators were adapting technology used to train World War II service personnel for the classroom, the name of the organization became the Department of Visual Instruction (DAVI). Twelve years later, DAVI became an affiliate of the NEA and finally the autonomous association, AECT, in 1974.

Today, AECT keeps an eye on the future of instructional technology while assisting educators with the changes and challenges that face them now. AECT members are professionals devoted to quality education. They care about doing their jobs better and want to embrace new methods, new equipment, and new techniques that assist learning.

AECT Affiliates

AECT has 45 state and 11 national affiliates, and has recently established several chapters surrounding major universities and metropolitan areas. These affiliated organizations add a localized dimension to your AECT membership and allow for more interaction among your colleagues. For more details on chapters and affiliates in your area, contact the AECT National Office.

AECT National Convention and INFOCOMM International Exposition

Each year, AECT brings top speakers to exciting locations, and presents over 300 sessions and special events to provide the best training available in the use of media in education and instruction. The convention features the INFOCOMM International Exposition, an extravaganza of over 300 companies showing the latest products to use in education and training.

Research and Theory Division (RTD) improves the design, execution, utilization, evaluation, and dissemination of audiovisual communications research; advises educators on the effective utilizations of the results of the research; and promotes applied and theoretical research on the use of media.

Membership Enrollment Form

YES! Enroll me for the membership category, division and options I have selected below. I understand that my membership will extend for 12 months from the date on this enrollment form.

CATEGORIES

☐ $50 **Regular Membership.** Fee includes $12 for *TechTrends* and (1) free division.

☐ $80 **Regular Comprehensive Membership.** Fee includes $12 for *TechTrends*, $30 for *Educational Technology Research and Development* and (1) free division.

☐ $20 **Student Membership.** Fee includes $12 for *TechTrends*, and (1) free division.
Name of University ____________ Head of Department ____________

☐ $40 **Student Comprehensive Membership.** Fee includes $12 for *TechTrends*, $20 for *Educational Technology Research and Development* and (1) free division.
Name of University ____________ Head of Department ____________
Note: Student memberships apply to students enrolled in formal coursework in the instructional technology field on at least a half-time basis.

☐ $62 International Regular Membership

☐ $100 International Comprehensive Membership

☐ $300 Corporate Membership

Division Memberships. Choose your one free division.

☐ DEMM ☐ ITED ☐ DISC ☐ DID ☐ INTL
☐ MDPD ☐ RTD ☐ DSMS ☐ DOT

OPTIONS

☐ $10 each Additional Division affiliation .... Total $______

☐ DEMM ☐ ITED ☐ DISC ☐ DID ☐ INTL
☐ MDPD ☐ RTD ☐ DSMS ☐ DOT

PAYMENT

Please indicate form of payment below. Remember, contributions or gifts to the Association for Educational Communications and Technology are not deductible as charitable contributions. However, they may be deductible as ordinary and necessary business expenses.

☐ My check or money order, payable to AECT, is enclosed for .... $______

☐ Please charge $______ to the following credit card account:
Master Card No. ____________ VISA Card No. ____________ Expires ____________
Signature ________________________________

☐ Please bill me.

Name ________________________________
Address ________________________________
City ____________ State ________ Zip ________ Daytime Telephone (____) ____________

Detach and mail to AECT, 1025 Vermont Avenue, N.W., Washington, DC 20005 or Fax to (202) 347-7839.

PREVIOUS PROCEEDINGS
ERIC ED NUMBERS

YEAR	LOCATION	ED NUMBER
1979	New Orleans	171329
1980	Denver	194061
1981	Philadephia	207487
1982	Dallas	223191 to 223236
1983	New Orleans	231337
1984	Dallas	243411
1985	Anaheim	256301
1986	Las Vegas	267753
1987	Atlanta	285518
1988	New Orleans	295621
1989	Dallas	308805
1990	Anaheim	323912

Table of Contents

Citation Networks of Selected Instructional Design and Technology Journals, 1985 - 1990
By Gary J. Anglin and Robert L. Towers 1

Mathematics Problem Solving within a Cooperative Learning Computer Structure
By Bernard W. Arenz 21

Effects of Variations in Learner Control on Children's Curiosity and Learning from Interactive Video
By Marilyn Plavocos Arnone and Barbara L. Grabowski 45

Global Navigation in the Year 2001
By Ronald Aust and Jim Klayder 69

The Interaction of Color Realism and Pictorial Recall Memory
By Louis H. Berry 83

Visual Complexity and Pictorial Memory: A Fifteen Year Research Perspective
By Louis H. Berry 91

Harnessing the Power of Interactivity for Instruction
By Terry K. Borsook 103

Doctoral Dissertation Research in Instructional Design and Technology from 1976 through 1988
By Edward P. Caffarella and Steven G. Sachs 119

The Facilitation of Vocabulary Acquisition by Embedding Explication in Expository Text
By Loretta A. Cardinale 131

The Effects of Program Control, Learner Control, and Learner Control with Advisement Lesson Control Strategies on Anxiety and Learning from Computer-assisted Instruction
By Randall P. Coorough 147

Using Photo-Interviewing as Tool for Research and Evaluation
By John V. Dempsey and Susan A. Tucker 167

Square One Television and Gender
By Ann De Vaney and Alejandra Elenes 191

A Comparison of an Expert Systems Approach to Computerized Adaptive Testing and an Item Response Theory Model
By Theodore W. Frick 213

Qualitative Research Methodology Issues
By Robert G. George 235

Supporting Learning with Process Tools: Theory and Design Issues
By David A. Goodrum and Randy A. Knuth 245

CBI: Systems or Medium?
By Nancy L. Higginbotham-Wheat 261

Learners Self-reports of Characteristics Related to Academic Achievement and Motivation
By Ann R. Igoe 275

The Effects of Pictorial Complexity and Cognitive Style on Visual Recall Memory
By Romaine R. Jesky and Louis H. Berry 289

Chaos, Communication, and Educational Technology
By James W. King 297

A Fine-tuned Look at White Space Variation in Desktop Publishing
By Nancy Nelson Knupfer and Marina Stock McIsaac 313

Educational Computing Social Foundations: A Symposium 329
- Educational Computing and Teachers: Changing Roles, Changing Pedagogy
  - By Nancy Nelson Knupfer
- Teachers, Computers, and Power
  - By Robert Muffoletto
- Social, Political, and Economic Factors that Affect Growth of Distance Education
  - By Marina Stock McIsaac
- The "History" of Technology and Education
  - By Alfred Bork
- Educational Technology Curriculum Theory: Toward a New Language of Possibility
  - By J. Randall Koetting

Visual Discrimination, Learning, and Page Layout: Wrap-around, Run-around, and Transparent Text
By Nancy Nelson Knupfer and Marina Stock McIsaac 383

Theory Building and Educational Technology: Foundations for Reconceptualization
By J. Randall Koetting and Alan Januszewski 395

The Effects of Feedback and Second Try in Computer-assisted Instruction for Rule-learning Task
By Doris Lee, Patricia Smith and Wilhelmina Savenye 411

Metacognitive and Cognitive Effects of Different Loci of Instructional Control
By Miheon Jo Lee 433

Implementation levels of a Videodisc-based Mathematics Program and Achievement
By William H. Lowry and Ron Thorkildsen 461

Learner Control versus Computer Control in Instructional Simulation
By Joseph S. Mattoon, James D. Klein and Richard A. Thurman 481

Interactive Media Audit Trails: Approaches and Issues
By Earl R. Misanchuk and Richard Schwier 499

Elementary Teacher Educational Media Utilization and Production in Selected Silicon Valley Schools
By John E. Morlan and Mei-Yan Lu 521

Interactive Video and Informal Learning Environments
By Kristine A. Morrissey 537

Learning Dynamic Processes from Animated Visuals in Microcomputer-based Instruction
By Gary B. Mayton 549

Instructional Analogies and the Learning of Tangible and Intangible Concepts
By Timothy J. Newby and Donald A. Stepich 561

A National Survey of Systemic School Restructuring Experiences
By Catherine A. Norris and Charles M. Reigeluth 581

Evaluation of Intelligent Tutoring Systems: What's been Done and What can be Done
By Michael A. Orey, Peter J. Legree, Philip D. Gillis and Estelle Bloom 593

Using Intelligent Tutoring Design Principles to Integrate Cognitive Theory into Computer-based Instruction
By Michael A. Orey and Wayne A. Nelson 611

Integrating Cognitive Theory into Gagne's Instructional Events
By Michael A. Orey, James R. Okey, Marshall G. Jones and Linda S. Stanley 625

Adult Attitudes toward Alternative Delivery Systems and Industrial Training Outcomes
By Rita C. Richey 635

The Effects of Visual Grouping on Learning from Computer Animated Presentations
By Lloyd P. Rieber 681

Computer-based Microworlds: A Bridge between Constructivism and Direct Instruction
By Lloyd P. Rieber 691

Adjunct Questions and Rule Learning
By Steven Rodriguez 709

An Evaluation of Alternative Distance Tutoring Models for At-Risk Elementary School Children
By Steven M. Ross, Gary R. Morrison, Lana J. Smith and Evelyn Cleveland 723

The Development of Writing Skills in a Computer-Intensive Environment
By Steven M. Ross, Lana J. Smith and Ellen Woodson 741

Receive-site Facilitator Practices and Student Performance in Satellite-delivered Instruction
By Floyd K. Russell 759

The Use of Instructional Development Procedures to Create Exhibits: A Survey of Major American Museums
By George W. Schlenk and Sharon A. Shrock 779

Distance Education in Taiwan: A Model Validated
By Mei-Yau Shih and Susan M. Zvacek 795

Effects on Concretely versus Abstractly Illustrated Instruction on Learning Abstract Concepts
By Margaret Ann Smith and Patricia L. Smith 803

The Validity of a Multiple-Choice, Paper and Pencil Instrument in Discriminating between Masters and Nonmasters of Instructional Design
By Sidney Leland Stepp and Sharon A. Shrock 817

Research on Hypermedia Browsers and Suggestions for Future Research
By William B. Story, Jr. and Francis A. Harvey 843

Instructional Design and Human Practice: What can we Learn from Habermas' Theory of Technical and Practical Human Interests?
By Michael J. Streibel 849

Automating Instructional Systems Development
By Robert D. Tennyson and Dean L. Christensen 885

Two Theories of Design and Instructional Design
By Steven D. Tripp 917

Color realism and Hemispheric Laterality
By Esther Waltz and Louis H. Berry 929

Instructor's Plan: A Lesson Planning Expert System for School Teachers
By Dennis A. Wilkins, Paul F. Cook and Edward E. Green 941

Cognitive Apprenticeships: An Instructional Design Review
By Brent Wilson and Peggy Cole 945

A Qualitative Study of Teachers' Planning of Instruction for Adult Learners in a Telecommunications-based Distance Education Environment
By Linda L. Wolcott 957

An Anthropological View of Educational Communications and Technology: Beliefs and Behaviors in Research and Theory
By Andrew R. J. Yeaman 979

Sociocultural Aspects of Computers in Education
By Andrew R. J. Yeaman 989

Survey of Instructional Design Needs and Competencies
By Barbara Seels and Z. Glasgow 995

1991

AUTHOR INDEX

Authors	Descriptors	YR PG
Anglin, G	instructional design	91/1
Anglin, G	networks	91/1
Arenz, B	cooperative learning	91/21
Arenz, B	problem solving	91/21
Arnone, M	interactive video	91/45
Arnone, M	learner control	91/45
Aust, R	navigation tools	91/69
Aust, R	networks	91/69
Berry, E	hemispheric laterality	91/929
Berry, E	realism	91/929
Berry, L	cognitive style	91/289
Berry, L	memory	91/83,91
Berry, L	realism	91/83
Berry, L	visuals	91/91
Berry, L	visual recall	91/289
Bloom, E	expert systems	91/593
Bloom, E	tutoring	91/593
Bork, A	education	91/349
Bork, A	technology	91/349
Borsook, T	interactivity	91/103
Borsook, T	instruction	91/103
Caffarella, E	instructional design	91/119
Caffarella, E	instructional technology	91/119
Cardinale, L	information processing	91/131
Cardinale, L	learning	91/131
Christensen, D	instructional development	91/885
Christensen, D	expert systems	91/885
Cleveland, E	telecommunications	91/723
Cleveland, E	tutoring	91/723
Cole, P	cognitive apprenticeship	91/945
Cole, P	instructional design	91/945
Cook, P	authoring	91/941
Cook, P	expert systems	91/941
Coorough, R	anxiety	91/147
Coorough, R	computer-assisted instruction	91/147
Dempsey, J	evaluation	91/167
Dempsey, J	qualitative research	91/167
De Vaney, A	gender	91/191
De Vaney, A	television	91/191
Elenes, A	gender	91/191
Elenes, A	television	91/191
Frick, T	expert systems	91/213
Frick, T	response theory	91/213
George, R	qualitative research	91/235
George, R	research methodology	91/235
Gillis, P	expert systems	91/593
Gillis, P	tutoring	91/593
Glasgow, Z	instructional design	91/995
Glasgow, Z	needs assessment	91/995
Goodrum, D	design	91/245
Goodrum, D	process tools	91/245
Grabowski, B	learner control	91/45
Grabowski, B	interactive video	91/45
Green, E	authoring	91/941

Authors	Descriptors	YR PG
Green, E	expert systems	91/941
Harvey, F	hypermedia	91/843
Harvey, F	research	91/843
Higginbotham-Wheat, N	computer-based instruction	91/261
Higginbotham-Wheat, N	systems	91/261
Igoe, A	academic achievement	91/275
Igoe, A	motivation	91/275
Januszewski, A	educational technology	91/395
Januszewski, A	theory	91/395
Jesky, R	cognitive style	91/289
Jesky, R	visual recall	91/289
Jones, M	cognitive theory	91/625
Jones, M	instructional strategies	91/625
King, J	communication	91/297
King, J	educational technology	91/297
Klayder, J	navigation tools	91/69
Klayder, J	networks	91/69
Klein, J	learner control	91/481
Klein, J	simulation	91/481
Knupfer, N	design	91/313,383
Knupfer, N	desktop publishing	91/313,383
Knupfer, N	educational computing	91/331
Knupfer, N	pedagogy	91/331
Knuth, R	design	91/245
Knuth, R	process tools	91/245
Koetting, J	educational technology	91/395
Koetting, J	theory	91/395
Koetting, R	curriculum theory	91/381
Koetting, R	educational technology	91/381
Lee, D	computer- assisted Instruction	91/411
Lee, D	feedback	91/411
Lee, M	loci of control	91/433
Lee, M	metacognition	91/433
Legree, P	expert systems	91/593
Legree, P	tutoring	91/593
Lowry, W	achievement	91/461
Lowry, W	videodisc	91/461
Lu, M	educational media	91/521
Lu, M	utilization	91/521
Mattoon, J	learner control	91/481
Mattoon, J	simulation	91/481
Mayton, G	animation	91/549
Mayton, G	computer-based instruction	91/549
McIsaac, M	design	91/313,383
McIsaac, M	desktop publishing	91/313,383
McIsaac, M	distance education	91/343
McIsaac, M	networks	91/343
Misanchuk, E	audit	91/499
Misanchuk, E	interactive media	91/499
Morlan, J	educational media	91/521
Morlan, J	utilization	91/521
Morrison, G	tutoring	91/723
Morrison, G	telecommunications	91/723
Morrissey, K	learning	91/537

Authors	Descriptors	YR PG
Morrissey, K	interactive video	91/537
Muffoletto, R	empowerment	91/341
Muffoletto, R	computers	91/341
Nelson, W	cognitive theory	91/611
Nelson, W	tutoring	91/611
Newby, T	concept development	91/561
Newby, T	learning	91/561
Norris, C	education	91/581
Norris, C	restructuring	91/581
Orey, M	cognitive theory	91/593,625
Orey, M	expert systems	91/593
Orey, M	instructional strategies	91/625
Orey, M	tutoring	91/593,611
Okey, J	cognitive theory	91/625
Okey, J	instructional strategies	91/625
Reigeluth, C	education	91/581
Reigeluth, C	restructuring	91/581
Richey, R	attitude	91/635
Richey, R	training	91/635
Rieber, L	animation	91/681
Rieber, L	constructionism	91/691
Rieber, L	microworlds	91/691
Rieber, L	visuals	91/681
Rodriguez, S	instructional strategies	91/709
Rodriguez, S	learning	91/709
Ross, S	computers	91/741
Ross, S	telecommunications	91/723
Ross, S	tutoring	91/723
Ross, S	writing	91/741
Russell, F	distance education	91/759
Russell, F	satellite-delivered instruction	91/759
Sachs, S	instructional design	91/119
Sachs, S	instructional technology	91/119
Savenye, W	computer- assisted instruction	91/411
Savenye, W	feedback	91/411
Schlenk, G	instructional development	91/779
Schlenk, G	research	91/779
Schwier, R	audit	91/499
Schwier, R	interactive media	91/499
Seels, B	instructional design	91/995
Seels, B	needs assessment	91/995
Shih, M	distance education	91/795
Shih, M	models	91/795
Shrock, S	evaluation	91/817
Shrock, S	instructional design	91/817
Shrock, S	instructional development	91/779
Shrock, S	research	91/779
Smith, L	computers	91/741
Smith, L	telecommunications	91/723
Smith, L	tutoring	91/723
Smith, L	writing	91/741
Smith, M	design	91/803
Smith, M	visuals	91/803
Smith, P	computer- assisted instruction	91/411

Authors	Descriptors	YR PG
Smith, P	design	91/803
Smith, P	feedback	91/411
Smith, P	visuals	91/803
Stanley, L	cognitive theory	91/625
Stanley, L	instructional strategies	91/625
Stepich, D	concept development	91/561
Stepich, D	learning	91/561
Stepp, S	evaluation	91/817
Stepp, S	instructional design	91/817
Story, W	hypermedia	91/843
Story, W	research	91/843
Streibel, M	instructional design	91/849
Streibel, M	theory	91/849
Tennyson, R	expert systems	91/885
Tennyson, R	instructional development	91/885
Thorkildsen, R	achievement	91/461
Thorkildsen, R	videodisc	91/461
Thurman, R	learner control	91/481
Thurman, R	simulation	91/481
Towers, R	instructional design	91/1
Towers, R	networks	91/1
Tripp, S	instructional design	91/917
Tripp, S	theory	91/917
Tucker, S	evaluation	91/167
Tucker, S	qualitative research	91/167
Waltz, E	hemispheric laterality	91/929
Waltz, E	realism	91/929
Wilkins, D	authoring	91/941
Wilkins, D	expert systems	91/941
Wilson, B	cognitive apprenticeship	91/945
Wilson, B	instructional design	91/945
Wolcott, L	distance education	91/957
Wolcott, L	teaching	91/957
Woodson, E	computers	91/741
Woodson, E	writing	91/741
Yeaman, A	computers	91/989
Yeaman, A	education	91/989
Yeaman, A	educational technology	91/979
Yeaman, A	theory	91/979
Zvacek, S	distance education	91/795
Zvacek, S	models	91/795

1991

DESCRIPTOR INDEX

Descriptors	Authors	YR PG
academic achievement	Igoe, A	91/275
achievement	Lowry, W	91/461
achievement	Thorkildsen, R	91/461
animation	Mayton, G	91/549
animation	Rieber, L	91/681
anxiety	Coorough, R	91/147
attitude	Richey, R	91/635
audit	Misanchuk, E	91/499
audit	Schwier, R	91/499
authoring	Cook, P	91/941
authoring	Green, E	91/941
authoring	Wilkins, D	91/941
cognitive apprenticeship	Cole, P	91/945
cognitive apprenticeship	Wilson, B	91/945
cognitive style	Berry, L	91/289
cognitive style	Jesky, R	91/289
cognitive theory	Jones, M	91/625
cognitive theory	Nelson, W	91/611
cognitive theory	Okey, J	91/625
cognitive theory	Orey, M	91/593,625
cognitive theory	Stanley, L	91/625
communication	King, J	91/297
computers	Muffoletto, R	91/341
computers	Ross, S	91/741
computers	Smith, L	91/741
computers	Woodson, E	91/741
computers	Yeaman, A	91/989
computer-assisted instruction	Coorough, R	91/147
computer- assisted instruction	Lee, D	91/411
computer- assisted instruction	Savenye, W	91/411
computer- assisted instruction	Smith, P	91/411
computer-based instruction	Higginbotham-Wheat, N	91/261
computer-based instruction	Mayton, G	91/549
concept development	Newby, T	91/561
concept development	Stepich, D	91/561
constructionism	Rieber, L	91/691
cooperative learning	Arenz, B	91/21
curriculum theory	Koetting, R	91/381
desktop publishing	Knupfer, N	91/313,383
desktop publishing	McIsaac, M	91/313,383
design	Goodrum, D	91/245
design	Knupfer, N	91/313,383
design	Knuth, R	91/245
design	McIsaac, M	91/313,383
design	Smith, M	91/803
design	Smith, P	91/803
distance education	McIsaac, M	91/343
distance education	Russell, F	91/759
distance education	Shih, M	91/795
distance education	Shih, M	91/795
distance education	Wolcott, L	91/957
distance education	Zvacek, S	91/795
education	Bork, A	91/349
education	Norris, C	91/581

Descriptors	Authors	YR PG
education	Reigeluth, C	91/581
education	Yeaman, A	91/989
educational computing	Knupfer, N	91/331
educational media	Morlan, J	91/521
educational media	Lu, M	91/521
educational technology	Januszewski, A	91/395
educational technology	Koetting, J	91/395
educational technology	Koetting, R	91/381
educational technology	King, J	91/297
educational technology	Yeaman, A	91/979
empowerment	Muffoletto, R	91/341
evaluation	Dempsey, J	91/167
evaluation	Shrock, S	91/817
evaluation	Stepp, S	91/817
evaluation	Tucker, S	91/167
expert systems	Bloom, E	91/593
expert systems	Christensen, D	91/885
expert systems	Cook, P	91/941
expert systems	Frick, T	91/213
expert systems	Green, E	91/941
expert systems	Gillis, P	91/593
expert systems	Legree, P	91/593
expert systems	Orey, M	91/593
expert systems	Tennyson, R	91/885
expert systems	Wilkins, D	91/941
feedback	Lee, D	91/411
feedback	Savenye, W	91/411
feedback	Smith, P	91/411
gender	De Vaney, A	91/191
gender	Elenes, A	91/191
hemispheric laterality	Berry, E	91/929
hemispheric laterality	Waltz, E	91/929
hypermedia	Harvey, F	91/843
hypermedia	Story, W	91/843
information processing	Cardinale, L	91/131
instruction	Borsook, T	91/103
instructional design	Anglin, G	91/1
nstructional design	Caffarella, E	91/119
instructional design	Cole, P	91/945
instructional design	Glasgow, Z	91/995
instructional design	Sachs, S	91/119
instructional design	Seels, B	91/995
instructional design	Shrock, S	91/817
instructional design	Stepp, S	91/817
instructional design	Streibel, M	91/849
instructional design	Towers, R	91/1
instructional design	Tripp, S	91/917
instructional design	Wilson, B	91/945
instructional development	Christensen, D	91/885
instructional development	Schlenk, G	91/779
instructional development	Shrock, S	91/779
instructional development	Tennyson, R	91/885
instructional strategies	Jones, M	91/625
instructional strategies	Okey, J	91/625

Descriptors	Authors	YR PG
instructional strategies	Orey, M	91/625
instructional strategies	Rodriguez, S	91/709
instructional strategies	Stanley, L	91/625
instructional technology	Caffarella, E	91/119
instructional technology	Sachs, S	91/119
interactivity	Borsook, T	91/103
interactive media	Misanchuk, E	91/499
interactive media	Schwier, R	91/499
interactive video	Arnone, M	91/45
interactive video	Grabowski, B	91/45
interactive video	Morrissey, K	91/537
learner control	Arnone, M	91/45
learner control	Grabowski, B	91/45
learner control	Klein, J	91/481
learner control	Mattoon, J	91/481
learner control	Thurman, R	91/481
learning	Cardinale, L	91/131
learning	Morrissey, K	91/527
learning	Newby, T	91/561
learning	Rodriguez, S	91/709
learning	Stepich, D	91/561
loci of control	Lee, M	91/433
memory	Berry, L	91/83,91
metacognition	Lee, M	91/433
microworlds	Rieber, L	91/691
models	Shih, M	91/795
models	Zvacek, S	91/795
motivation	Igoe, A	91/275
navigation tools	Aust, R	91/69
navigation tools	Klayder, J	91/69
needs assessment	Glasgow, Z	91/995
needs assessment	Seels, B	91/995
networks	Anglin, G	91/1
networks	Aust, R	91/69
networks	Klayder, J	91/69
networks	McIsaac, M	91/343
networks	Towers, R	91/1
pedagogy	Knupfer, N	91/331
problem solving	Arenz, B	91/21
process tools	Goodrum, D	91/245
process tools	Knuth, R	91/245
qualitative research	Dempsey, J	91/167
qualitative research	George, R	91/235
qualitative research	Tucker, S	91/167
realism	Berry, E	91/929
realism	Berry, L	91/83
realism	Waltz, E	91/929
research	Harvey, F	91/843
research	Schlenk, G	91/779
research	Shrock, S	91/779
research	Story, W	91/843
research methodology	George, R	91/235
response theory	Frick, T	91/941
restructuring	Norris, C	91/581

Descriptors	Authors	YR PG
restructuring	Reigeluth, C	91/581
satellite-delivered instruction	Russell, F	91/759
simulation	Klein, J	91/481
simulation	Mattoon, J	91/481
simulation	Thurman, R	91/481
systems	Higginbotham-Wheat, N	91/261
teaching	Wolcott, L	91/957
technology	Bork, A	91/349
telecommunications	Cleveland, E	91/723
telecommunications	Morrison, G	91/723
telecommunications	Ross, S	91/723
telecommunications	Smith, L	91/723
television	De Vaney, A	91/191
television	Elenes, A	91/191
theory	Januszewski, A	91/395
theory	Koetting, J	91/395
theory	Streibel, M	91/849
theory	Tripp, S	91/917
theory	Yeaman, A	91/979
training	Richey, R	91/635
tutoring	Bloom, E	91/593
tutoring	Cleveland, E	91/723
tutoring	Gillis, P	91/593
tutoring	Legree, P	91/593
tutoring	Morrison, G	91/723
tutoring	Nelson, W	91/611
tutoring	Orey, M	91/593,611
tutoring	Ross, S	91/723
tutoring	Smith, L	91/723
utilization	Lu, M	91/521
utilization	Morlan, J	91/521
videodisc	Lowry, W	91/461
videodisc	Thorkildsen, R	91/461
visuals	Berry, L	91/91
visuals	Rieber, L	91/681
visuals	Smith, M	91/803
visuals	Smith, P	91/803
visual recall	Berry, L	91/289
visual recall	Jesky, R	91/289
writing	Ross, S	91/741
writing	Smith, L	91/741
writing	Woodson, E	91/741

Title:

Citation Networks of Selected Instructional Design and Technology Journals, 1985 - 1990

Authors:

Gary J. Anglin
Robert L. Towers

Citation Networks of Selected Instructional Design and Technology Journals, 1985-1990

The field of instructional design and technology is certainly at a stage of development where it's history, current status, and future deserve attention by researchers in the field. So (1988) suggests that "Self-reflection is a mark of maturity" (So, 1988; p. 236). One approach used for exploring the patterns of communication in a discipline is bibliometric research. It is assumed results from such research will provide one indicator of the structure of the knowledge base in a field, as well as the disciplinary boundaries (e.g., invisible colleges within a discipline, Crane, 1972). Garfield (1979) has pointed out potential problems with citation studies, including the lack of quality indicators of the particular scholarly work, intentions for citing authors is not made clear, potential bias for recent publications, and the fact that all citations in many studies are weighted equally. After reviewing studies which have addressed some of Garfield's concerns, So (1988), concludes that "it seems fair to conclude that citation analysis is a useful method if it is employed carefully" (p. 237).

Several investigators have attempted to reflect on the nature and structure of the field of instructional design and technology (Braden, 1981; Dick, & Dick, 1989; Durzo, Diamond, & Doughty, 1979; Gentry & Csete, 1991; Reigelith, 1989; Sachs, 1984; Heinich (1984); Wedman, 1987; Winn, 1989). A majority of these papers are theoretical or conceptual in nature while two are bibliographic studies.

The primary purpose of this paper is to update and extend the study conducted by Sachs (1984). Using citations from two journals, Journal of Instructional Development (JID), (Spring, 1980 - Winter, 1983) and Performance and Instruction Journal (P&I), (June, 1980 - May, 1983), the study by Sachs (1984) addressed questions concerning patterns of citation, frequency of cited authors, identification of invisible colleges, and systematic literature development. Sachs concluded that there are invisible colleges within the field of instructional development, as a whole there is poor linkage in the field, and that across the two journals included, the groups identified and the linkages are different. Sachs concluded that "based on the findings from this study, the literature in JID and P&I are reporting more of the art and less of the science of instructional development." (p. 13).

The current study will address the following three questions:

1. Who are the most cited authors in the field?
2. Do invisible colleges exist in the field of instructional design and technology?
3. If invisible college exist in the field, who are the participants in each invisible colleg

Method

Data Source

The present study will include the Journal Instructional Development (JID, Spring 1985 - Fall 1990), Educational Communication and Technology Journal (ECTJ, Spring 1985 - Summer 1990), and Performance Improvement Quarterly (PIQ, Spring 1988 - Fall 1990). It should be noted that JID and ECTJ are now printed under one cover (Educational Technology Research and Development (ETRD)). However, the identity of each journal (JID and ECTJ) is preserved including separate editors and editorial boards. We selected PIQ, a relatively new journal rather that P&I since in our judgement PIQ includes more theoretical papers and research studies. Both journals are published by the same national organization.

We have also extended the data set by increasing the number of years for journal inclusion to five.

Procedure

For Each of the three journals included in the study (ECTD, JID, PIQ) reference lists included in each article or review were collected. The name of each author, co-author or editor was entered in a data base. Information entered for each author included name, date of citation, volume and issue number of the journal. An alphabetical sort was then performed by author. Using the sorted data base listing, the number of citations per author was recorded.

Individuals were included in the study if they were cited a minimum of 5 times (after Sachs, 1984). A matrix was then constructed which identified the number of citations for each article by author. This matrix was required for the subsequent analysis.

Results and Discussion

Descriptive Information

The total number of authors cited in all reference lists collected as well as the number of authors included in the study for each journal is reported in Table 1. The total number of citations for frequently cited authors by journal are listed in Tables 2, 3, and 4.

Insert Tables 1,2, 3, and 4

From the 12,220 citations entered in the data base for all three journals, a total of 386 individuals were selected, 33 from PIQ, 131 from JID, and 222 from ECTJ (See Table 1.). As indicated in Table 1, an individual was included if they received 5 or more citations for the journal under consideration.

The highest numbers of citations reported were 83 (R. M. Gagne), 76 (R. D. Tennyson), and 43 (R. Kaufman) for JID, ECTJ, and PIQ, respectively.

Hierarchical Cluster Analysis

The purpose of the hierarchical cluster analysis was to determine if the data could be arranged into homogeneous groups. The data was analyzed using the Johnson (1967) maximum (compactness) method. Results of the cluster analysis are reported in Tables 5, 6, and 7.

Insert Tables 5, 6, and 7

The clusters identified among frequently cited individuals suggest that homogeneous groups do exist. There were 21, 25, and 7 homogeneous groups identified for ECTJ/ETRD(Research), JID/ETRD(Development), and PIQ respectively. It should be noted that PIQ is a relatively new journal and the number of data points included is much smaller than for the other journals included.

The cluster analysis yielded 5 clusters with 8 or more frequently cited individuals for JID. For JID, 16 groups of highly cited individuals are relatively small. There were 9 relatively large clusters identified for ECTJ. (Clusters 1.0 - 9.0 inclusive). The number of authors in each of the 9 large clusters ranged from 10 to 32. With the exception of cluster 1.0, the analysis for PIQ yielded seven small homogeneous groups of individuals. This was expected since the journal began publishing in 1988.

For many of the groups dominant individuals (highly cited) can be identified. For example, R. M. Gagne was cited 83 times in JID (See Table 5, cluster 1.0). The the results of the study support the conclusion that there are "many" invisible colleges in the field. This may be the case since the field is quite diverse. It is also a possible indication that the field is highly fragmented and not clearly defined. However, the groups of frequently cited individuals do significantly influence the development of the field and the practice of instructional design and development.

References

Braden, R. A. (1981). One hundred book titles: A twelve-foot shelf of basic references for instructional design and development. Educational Technology, 21(9), 41-45.

Crane, D. (1972). Invisible colleges: Diffusion of knowledge in scientific communities. Chicago: The University of Chicago Press.

Dick, W., & Dick, W. D. (1989). Analytical and empirical comparisons of the Journal of Instructional Development and Educational Communication and Technology Journal. Educational Technology Research and Development, 37, 81-87.

Durzo, J. J., Diamond, R. M., & Doughty, P. L. (1979). Selected references related to the practice of instructional development: A place to start. Journal of Instructional Development, 2(4), 19-22.

Garfield, E. (1979. Perspective on citation analysis of scientists. In Citation indexing, (pp. 240-252). Philadelphia: ISI Press.

Gentry, C., & Csete, J. (1991). Educational technology in the 1990s. In G. J. Anglin (Ed.), Instructional technology: Past, present and future (pp. 20-33). Englewood, CO: Libraries Unlimited.

Heinich, R. (1984). The proper study of instructional technology. Educational Communication and Technology Journal, 32(2), 67-87.

Johnson, S. C. (1967). Hierarchical clustering schemes. Psychometrika, 32(3), 241-254.

Reigeluth, C. M. (1989). Educational technology at the crossroads: New mindsets and new directions. Educational Technology Research and Development, 37, 67-80.

Sachs, S. G. (1984). Citation patterns in instructional development literature. Journal of Instructional Development, 7(2), 8-13.

So, Clement Y. J.(1988). Citation patterns of core communication journals: An assessment of the developmental status of communication. Human Communication Research, 15(2), 236-255.

Wedman, J. F. (1987). Citation patterns in the computer-based instruction literature. Journal of Computer-Based Instruction, 14(3), 91-95.

Winn, W. (1989). Toward a rationale and theoretical basis for educational technology. Educational Technology Research and Development, 37, 35-46.

Table 1.0
Total Citations Identified in Selected Instructional Journals

Journal	Volumes Used in Study	Total Citations	Citations Included*
PIQ	Spring 1988 - Fall 1990	1805	33
JID/ETRD (Development Section)	Spring 1985 - Fall 1990	3978	131
ECTJ/ETRD (Research Section)	Spring 1985 - Summer 1990	6437	222
	Totals	12,220	

* Individuals with five or more citations included in the study.

TABLE 2.0

Number of Citations for Frequently Cited Authors

JID/ETRD (Development Section)

Author	Citations	Author	Citations	Author	Citations
Gagne, R.M.	83	Feightner, J.W.	8	Wildman, T.M.	6
Reigeluth, C.M.	79	Hannum, W.H.	8	Wittrock, M.C.	6
Briggs, L.J.	49	Morgan, R.M.	8	Alesandrini, K.L.	5
Jonassen, D.H.	42	O'Neil, H.F. Jr.	8	Alessi, S.M.	5
Hannafin, M.J.	32	Readence, J.E.	8	Bandura, A.	5
Merrill, M.D.	26	Romiszowski, A.J.	8	Bean, T.W.	5
Keller, J.M	26	Sleeman, D.H.	8	Becker, H.J.	5
Dick, W.	25	Driscoll, M.P.	7	Berman, P.	5
Tennyson, R.D.	23	Grooper, G.L.	7	Bork, A.	5
Bratton, B.	20	Hooper, K.	7	Boyer, E.L.	5
Reiser, R.A.	19	Hooper, S.	7	Burkman, E.	5
Clark, R.E.	18	Kincaid, J.P.	7	Christensen, D.L.	5
Glaser, R.	17	McCombs B.L.	7	Dalton, D.W.	5
Kaufman, R.A.	17	Morrison, G.R.	7	Davis, R.H.	5
Kearsley, G.P.	17	Novak, J.D.	7	Daynes, R.	5
Norman, G.R.	16	Rossett, A.	7	Deden-Parker, A.	5
Sullivan, H.J.	16	Salomon, G.	7	Ericsson, K.A.	5
Anderson, J.R.	15	Schiffman, S.	7	Flavell,, J.H.	5
Branson, R.K.	15	Sprio, R.J.	7	Frase, L.T.	5
Rieber, L.	15	Wilson, R.G.	7	Gagne, E.	5
Brown, J.S.	14	Alen, V.L.	6	Gilbert, T.F.	5
Ross, S.M.	14	Allen, B.S.	6	Heyneman, S.	5
Goodson, L.A.	13	Barrows, H.S.	6	Hunter, M.	5
Andrews, D.H.	12	Bunderson, C.V.	6	Kinzie, M.B.	5
Carey, L.	12	Dansereau, D.E.	6	Lesgold, A.M.	5
Simon, H.A.	12	Davies, I.K.	6	Malon, T.W.	5
Wager, W.W.	12	DeBloois, M.L.	6	Markle, S.M.	5
Brown, A.L.	11	Foshay, W.R.	6	Martin, B.L.	5
Anderson, R.C.	10	Harmon, P.	6	Masia, B.B.	5
Bower, G.H.	10	Heinich, R.	6	Mayer, R.E.	5
Higgins, N.	10	Joyce, B.	6	Naisbitt, J.	5
Park, O.	10	Knirk, F.G.	6	Neufeld, V.R.	5
Snow, S.E.	10	Kulik, F.G.	6	Phillips, T.L.	5
Barby, B.	9	Levie, W.H.	6	Resnick, L.B.	5
Collins, A.M.	9	Levin, J.R.	6	Richards, B.F.	5
Craik, F.I.M.	9	Mager, R.F.	6	Rogers, E.M.	5
Gustafson, K.L.	9	McLaughlin, M.W.	6	Rumelhart, D.E.	5
Merrill, P.F.	9	Montague, W.E.	6	Scandura, J.M.	5
Misanchuk, E.R.	9	Norman, D.A.	6	Seidel, R.J.	5
Salisbury, D.F.	9	Pressley, M.	6	Shrock, S.A.	5
Shneiderman, B.	9	Trollip, S.R.	6	Silber, K.H.	5
Tessmer, M.A.	9	Tulving , E.	6	Snelbecker, G.E.	5
Ambron, S.R.	8	Walberg, H.J.	6	Stein, F.S.	5
Bloom, B.S.	8	Weinstein, C.E.	6		

TABLE 3.0

Number of Citations for Frequently Cited Authors

ECTJ/ETRD (Reasearch Section)

Author	Citations	Author	Citations	Author	Citations
Tennyson, R.D.	76	Piaget, J.	13	Lesgold, A.M.	9
Levin, J.R.	72	Wright, J.C.	13	Mastropieri, M.A.	9
Reigeluth, C.M.	62	Greeno, J.G.	12	Paivio, A.	9
Ross, S.M.	59	Jonassen, D.H.	12	Papert, S.	9
Clark, R.E.	59	Levie, W.H.	12	Scruggs, T.E.	9
Salomon, G.	54	O'Dell, J.K.	12	Wager, W.W.	9
Hannafin, M.J.	48	Reiser, R.A.	12	Wang, M.C.	9
Gagne, R.M.	48	Schramm, W.L.	12	Wartella, E.	9
Presley, M.	44	Slavin, R.E.	12	Watkins, B.A.	9
Kulik, J.A.	33	Bower, G.H.	11	Bandura, A.	8
Merrill, M.D.	32	Brown, A.L.	11	Belmont, J.M.	8
Mayer, R.E.	30	Davidson, G.V.	11	Bryant, J.	8
Snow, R.E.	28	Holyoak, K.J.	11	Burton, R.R.	8
Anderson, J.R.	28	McCormick, C.B.	11	Campbell, D.T.	8
Anderson, D.R.	28	Miller, G.E.	11	Fleming, M.L.	8
Glaser, R.	27	Rothen, W.	11	Huston, A.C.	8
Brown, J.S.	22	Webb, N.M.	11	Keller, F.S.	8
Carrier, C.A.	21	Williams, G.W.	11	Landa, L.N.	8
Morrison, G.R.	21	Atkinson, R.C.	10	Lincoln, Y.S.	8
Simon, H.A.	21	Bork, A.	10	Meurrens, M.	8
Heinich, R.	20	Branson, R.K.	10	O'Neil, H.F.	8
Briggs, L.J.	19	Christensen, D.L.	10	Palmer, E.L.	8
Sulivan, H.J.	19	Cohen, P.A.	10	Pellegrino, J.W.	8
Anderson, R.C.	17	Dick, W.	10	Resnick, L.B.	8
McCombs, B.L.	17	Gardner, H.	10	Romiszowski, A.	8
Butterfield, E.C.	16	Montague, W.E.	10	Simonson, M.R.	8
Rakow, E.A.	16	Peterson, P.L.	10	Stevens, A.L.	8
Williams, M.D.	16	Pettersson, R.	10	Winn, W.D.	8
Bransford, J.D.	15	Phillips, T.L.	10	Ball, S.A.	7
Chadwick, C.B.	15	Reder, L.M.	10	Bogatz, G.A.	7
Collins, A.M.	15	Suppes, P.	10	Carey, L.	7
Guba, E.G.	15	Anand, P.G.	9	Case, R.P.	7
Sleeman, D.H.	15	Ausubel, D.P.	9	Cook, T.D.	7
Wittrock, M.C.	15	Bloom, B.S.	9	Craik, F.I.M.	7
Ershov, A.P.	14	Bruner, J.S.	9	DeBloois, M.L.	7
Hartley, J.	14	Chipman, S.F.	9	Deci, E.L.	7
Lepper, M.R.	14	Collins, W.A.	9	Eisner, E.W.	7
Levin, S.R.	14	Cormier, S.M.	9	Flavell, J.H.	7
Lorch, E.P.	14	Cronbach, L.J.	9	Johnson, D.W.	7
Park, O.	14	Dalton, D.W.	9	Maehr, M.L.	7
Steinberg, R.J.	14	Garhart, C.	9	Malone, T.W.	7
Bangert, R.L.	13	Hagman, J.D.	9	McGivern, J.E.	7
Kulik, C.C.	13	Kerr, S.T.	9	Mielke, K.W.	7
Norman, D.A.	13	Kinzie, M.B.	9	Mosley, M.L.	7

TABLE 3.0 (Continued)
Number of Citations for Frequently Cited Authors
ETCJ/ETRD (Research Section)

Author	Citations	Author	Citations	Author	Citations
Newell, A.	7	Morles, A.	6	Krisak, N.	5
O'Shea, T.	7	Petkovich, M.D.	6	Liben, L.	5
Park, S.I.	7	Roberts, T.	6	Lockhart, K.A.	5
Reese, H.W.	7	Schon, D.A.	6	Markle, S.M.	5
Reif, F.	7	Sherman, J.G.	6	McDaniel, M.A.	5
Seidel, R.J.	7	Shulman, L.S.	6	Megarry, J.	5
Spiro, R.J.	7	Siegler, R.S.	6	Morgan, R.M.	5
Steinberg, E.R.	7	Stanley, J.C.	6	Nelson, T.O.	5
Streibel, M.J.	7	Windham, D.M.	6	Nunan, T.	5
Taylor, R.	7	Alwitt, L.F.	5	Pea, R.D.	5
Allessi, S.M.	6	Andrews, D.H.	5	Pezdek, K.	5
Bereiter, C.	6	Bangert-Downs, R.	5	Pittman, T.S.	5
Calvert, S.L.	6	Bryant, S.L.	5	Reed, S.K.	5
Carroll, J.M.	6	Bunderson, C.V.	5	Reeves, B.	5
Chi, M.T.H.	6	Buttrey, T.	5	Reiber, L.P.	5
Delaney, H.D.	6	Campione, J.C.	5	Rumelhart, D.E.	5
DiVesta, F.J.	6	Chen, M.	5	Scanlon, E.	5
Federico, P.A.	6	Dunn, T.G.	5	Schmid, R.F.	5
Fennema, E.	6	Dwyer, F.M.	5	Segal, J.W.	5
Franks, J.J.	6	Gage, N.L.	5	Sharan, S.	5
Gay, G.	6	Gentner, D.	5	Shriberg, L.K.	5
Gerlach, V.S.	6	Ghatala, E.S.	5	Stein, B.S.	5
Gibson, J.J.	6	Gick, M.L.	5	Thorndike, E.L.	5
Glass, G.V.	6	Goodson, L.A.	5	Tobias, S.	5
Jamison, D.	6	Greene, M.	5	Trollip, S.R.	5
Johnson, R.T.	6	Grooper, G.L.	5	Tulis, T.S.	5
Keller, J.M.	6	Guilford, J.P.	5	Waller, R.	5
Lentz, R.	6	Houghton, H.A.	5	Weil, M.	5
McCormick, D.B.	6	Johnson, C.W.	5	Willows, D.M.	5
Morgan, G.	6	Joyce, B.	5	Wilson, D.M.	5

TABLE 4.0

Number of Citations for Frequently Cited Authors

PIQ

Author	Citations	Author	Citations	Author	Citations
Kaufman, R.	43	Jay, S.J.	9	Langley, P.C.	6
Gagne, R.M.	18	Zemke, R.	9	Miller, P.L.	6
Rothwell, W.	16	Bloom, B.S.	7	Ambron, S.	5
Jonassen, D.H.	16	Gilbert, T.F.	7	Binder, C.V.	5
Gerlach, V.S.	15	Harless, J.	7	Campbell, D.T.	5
Glaser, R.	14	Hooper, K.	7	Clark, R.E.	5
Reigeluth, C.M.	13	Kirkpatrick, D.A.	7	Hartley, J.	5
Simon, H.A.	12	Patton, M.Q.	7	Knowles, M.S.	5
Smith, M.E.	11	Schmid R.F.	7	Landa, L.N.	5
Swanson, R.A.	10	Sobel, I.	7	Mager, R.F.	5
Anderson, J.G.	9	Wright, P.C.	7	Merrill, M.D.	5

TABLE 5.0

Hierarchical Clusters Suggesting Jointly Occupied Positions in the Network of Highly Cited Authors.

JID/ETRD (Development Section)

Author	Citations	Author	Citations	Author	Citations
1.0		**4.0**		**9.0**	
Craik, F.I.M.	9	Bork, A.	5	Burkman, E.	5
Flavel, J.H.	5	Levie, W.H.	6	Martin, B.L.	5
Scandura, J.M.	5	Allen, B.S.	6	Schiffman, S.	7
Montague, W.E.	6	Hooper, S.	7	Shrock, S.A.	5
Norman, D.A.	6	Morrison, G.R.	7	Tessmer, M.A.	9
Rumelhart, D.E.	5	Hannafin, M.J.	32	Wager, W.W.	12
Novak, J.D.	7	Ross, S.M.	14		
Stein, F.S.	5	Salisbury, D.F.	9	**10.0**	
Keller, J.M.	26			Ambron, S.R.	8
Merrill, M.D.	26	**5.0**		Hooper, K.	7
Tennyson, R.D.	23	Davis, R.H.	5	Malone, T.W.	5
Clark, R.E.	18	Anderson, J.R.	15	Shneiderman, B.	9
Glaser, R.	17	Collins, A.M.	9	Snow, R.E.	10
Jonassen, D.H.	42	Levin, J.R.	6		
Gagne, R.M.	83	Pressley, M.	6	**11.0**	
Briggs, L.J.	49	Seidel, R.J.	5	Kinzie, M.B.	5
Reigeluth, C.M.	79	Brown, J.S.	14	Davies, I.K.	6
Gagne, E.	5	Sleeman, D.H.	8	Markle, S.M.	5
Wildman, T.M.	6			Kaufman, R.A.	17
Gustafson, K.L.	9	**6.0**		Morgan, R.M.	8
Knirk, F.G.	6	Bunderson, C.V.	6		
Salomon, G.	7	Hunter, M.	5	**12.0**	
Wittrock, M.C.	6	Joyce, B.	6	Heinich, R.	6
		Christensen, D.L.	5	Bratton, B.	20
2.0		Park, O.	10	Reiser, R.A.	19
Carey, L.	12	Richards, B.F.	5	Silber, K.H.	5
Dick, W.	25	Tulving, E.	6		
Andrews, D.H.	12			**13.0**	
Goodson, L.A.	13	**7.0**		Rogers, E.M.	5
Ericsson, K.A.	5	Foshay, W.R.	6	Branson, R.K.	15
Simon, H.A.	12	Bloom, B.S.	8	Hannum, W.H.	8
Brown, A.L.	11	Masia, B.B.	5	Romiszowski, A.J.	8
Bower, G.H.	10	Grooper, G.L.	7		
Dansereau, D.E.	6	Gilbert, T.F.	5	**14.0**	
O'Neil, H.F. Jr.	8	Merrill, P.F.	9	Barrows, H.S.	6
Weinstein, C.E.	6	Resnick, L.B.	5	Feightner, J.W.	8
				Neufeld, V.R.	5
3.0		**8.0**		Norman, G.R.	16
Frase, L.T.	5	Driscoll, M.	7		
Rieber, L.	15	Snelbecker, G.E.	5	**15.0**	
Mayer, R.E.	5	Higgins, N.	10	Berman, P.	5
Anderson, R.C.	10	Sullivan, H.J.	16	McLaughlin, M.W.	6
Sprio, R.J.	7	Rossett, H.J.	7	Boyer, E.L.	5
Dalton, D.W.	5	Walberg, H.J.	6	Naisbitt, J.	5
Phillips, T.L.	5				
Alessi, S.M.	5				
Trollip, S.R.	6				

TABLE 5.0 (Continued)
Hierarchical Clusters Suggesting Jointly Occupied Positions in the Network of Highly Cited Authors.
JID/ETRD (Development Section)

Author	Citations	Author	Citations	Author	Citations
16.0		**17.0**		**18.0**	
Daynes, R.	5	Bandura, A.	5	Bean, T.W.	5
DeBloois, M.L.	6	Mager, R.F.	6	Readence, J.E.	8
Lesgold, A.M.	5	McCombs, B.L.	7		
				*Alesandrini, K.L.	5
19.0		**20.0**		*Becker, H.J.	5
Harmon, P.	6	Barby, R.	9	*Deden-Parker, A.	5
Kearsley, G.P.	17	Kincaid, J.P.	7	*Heyneman, S.	5
Wilson, R.G.	7			*Misanchuk, E.R.	9
21.0					
Alen, V.L.	6				
Kulik, F.G.	6				

* **Isolates**

TABLE 6.0

Hierarchical Clusters Suggesting Jointly Occupied Positions in the Network of Highly Cited Authors.

ECTJ/ETRD (Reaearch Section)

Author	Citations	Author	Citations	Author	Citations
1.0		**2.0**		**3.0**	
Campione, J.C.	5	Fleming, M.L.	8	Johnson, C.W.	5
Belmont, J.M.	8	Simonson, M.R.	8	Atkinson, R.C.	10
Butterfield, E.C.	16	Schramm, W.L.	12	Paivio, A.	9
Ghatala, E.S.	5	Winn, W.D.	8	McDaniel, M.A.	5
Cormier, S.M.	9	Eisner, E.W.	7	Bryant, S.L.	5
Gentner, D.	5	Joyce, B.	5	Miller, G.E.	11
Gick, M.L.	5	Schon, D.A.	6	Levin, J.R.	72
Hagman, J.D.	9	Weil, M.	5	Presley, M.	44
Nelson, T.O.	5	Landa, L.N.	8	Reese, H.W.	7
Newell, A.	7	Reiser, R.A.	12	Mastropieri, M.A.	9
Reif, F.	7	Heinich, R.	20	Delaney, H.D.	6
Brown, A.L.	11	Wittrock, M.C.	15	McCormick, C.B.	11
Flavell, J.H.	7	Ball, S.A.	7	McGivern, J.E.	7
Siegler, R.S.	6	Bogatz, G.A.	7	Scruggs, T.E.	9
Stein, B.S.	5	Cook, T.D.	7	Shriberg, L.K.	5
Burton, R.R.	8	Palmer, E.L.	8		
Brown, J.S.	22	Alwitt, L.F.	5	**5.0**	
Sleeman, D.H.	15	Huston, A.C.	8	Pellegrino, J.W.	8
Anderson, J.R.	28	Pezdek, K.	5	Carey, L.	7
Glaser, R.	27	Wartella, E.	9	Dick, W.	10
Greeno, J.G.	12	Anderson, D.R.	28	Andrews, D.H.	5
Simon, H.A.	21	Levin, S.R.	14	Goodson, L.A.	5
Snow, R.E.	28	Lorch, E.P.	14	Resnick, L.B.	8
Salomon, G.	54	Calvert, S.L.	6	Merrill, M.D.	32
Tennyson, R.D.	76	Watkins, B.A.	9	Gagne, R.M.	48
Clark, R.E.	59	Wright, J.C.	13	Reigeluth, C.M.	62
Franks, J.J.	6			Briggs, L.J.	19
Holyoak, K.J.	11	**4.0**		Wager, W.W.	9
Bransford, J.D.	15	Bunderson, C.V.	5	Lockhart, K.A.	5
Peterson, P.L.	10	Slavin, R.E.	12	McCombs, B.L.	17
Steinberg, R.J.	7	Campbell, D.T.	8	Webb, N.M.	11
Thorndike, E.L.	5	Stanley, J.C.	6		
		Seidel, R.J.	7	**6.0**	
7.0		Gay, G.	6	DiVesta, F.J.	6
Mosley, M.L.	7	Steinberg, E.R.	14	Gardner, H.	10
Maehr, M.L.	7	Hannafin, M.J.	48	Anand, P.G.	9
Sulivan, H.J.	19	Morrison, G.R.	21	Jonassen, D.H.	12
Jamison, D.	6	O'Dell, J.K.	12	Petkovich, M.D.	6
Suppes, P.	10	Davidson, G.V.	11	Reder, L.M.	10
Glass, G.V.	6	Rakow, E.A.	16	Romiszowski, A.	8
Cohen, P.A.	10	Ross, S.M.	59	Hartley, J.	14
Kulik, J.A.	33	Carrier, C.A.	21	Allessi, S.M.	6
Bangert, R.L.	13	Williams, M.D.	16	Trollip, S.R.	5
Kulik, C.C.	13			Bork, A.	10
Williams, G.W.	11			Tulis, T.S.	5

TABLE 6.0 (Continued)
Hierarchical Clusters Suggesting Jointly Occupied Positions in the Network of Highly Cited Authors.
ECTJ/ETRD (Research Section)

Author	Citations	Author	Citations	Author	Citations
8.0		**9.0**		**10.0**	
Cronbach, L.J.	9	Waller, R.	5	Bereiter, C.	6
Megarry, J.	5	Bandura, A.	8	Case, R.P.	7
Greene, M.	5	Keller, J.M.	6	Federico, P.A.	6
Pea, R.D.	5	Krisak, N.	5	Montague, W.E.	10
O'Shea, T.	7	McCormick, D.B.	6	Collins, A.M.	15
Bruner, J.S.	9	Kinzie, M.B.	9	Stevens, A.L.	8
Piaget, J.	13	Deci, E.L.	7		
O'Neil, H.F.	8	Lepper, M.R.	14	**13.0**	
Nunan, T.	5	Malone, T.W.	7	Anderson, R.C.	17
Streibel, M.J.	7	Wang, M.C.	9	Bower, G.H.	11
Papert, S.	9			Chi, M.T.H.	6
Taylor, R.	7	**12.0**		Ausubel, D.P.	9
		Craik, F.I.M.	7	Mayer, R.E.	30
11.0		Segal, J.W.	5		
Collins, W.A.	9	Norman, D.A.	13	**14.0**	
Bryant, J.	8	Rumelhart, D.E.	5	Lesgold, A.M.	9
DeBloois, M.L.	7	Spiro, R.J.	7	Lentz, R.	6
Dalton, D.W.	9			Levie, W.H.	12
Garhart, C.	9	**16.0**		Houghton, H.A.	5
Phillips, T.L.	10	Buttrey, T.	5	Willows, D.M.	5
		Park, O.	14		
15.0		Rothen, W.	11		
Dwyer, F.M.	5	Tobias, S.	5	**17.0**	
Grooper, G.L.	5			Liben, L.	5
Christensen, D.L.	10	**18.0**		Bangert-Downs, R.	5
Park, S.I.	7	Gerlach, V.S.	6	Guba, E.G.	15
		Markle, S.M.	5	Lincoln, Y.S.	8
19.0		Schmid, R.F.	5		
Chen, M.	5			**21.0**	
Reeves, B.	5	**20.0**		Bloom, B.S.	9
		Gibson, J.J.	6	Sherman, J.G.	6
22.0		Pettersson, R.	10	Keller, F.S.	8
Branson, R.K.	10			Shulman, L.S.	6
Carroll, J.M.	6	**23.0**			
		Chadwick, C.B.	15	**24.0**	
25.0		Morgan, R.M.	5	Johnson, D.W.	7
Ershov, A.P.	14	Windham, D.M.	6	Johnson, R.T.	6
Kerr, S.T.	9			Sharan, S.	5
*Dunn, T.G.	5	*Guilford, J.P.	5	*Meurrens, M.	8
*Morgan, G.	6	*Morles, A.	6	*Pittman, T.S.	5
*Reiber, L.P.	5	*Roberts, T.	6	*Scanlon, E.	5

* **Isolates**

TABLE 7.0

Hierarchical Clusters Suggesting Jointly Occupied Positions in the Network of Highly Cited Authors.

PIQ

Author	Citations	Author	Citations	Author	Citations
1.0		**2.0**		**3.0**	
Gilbert, T.F.	7	Bloom, B.S.	7	Binder, C.V.	5
Harless, J.	7	Gagne, R.M.	18	Clark, R.E.	5
Smith, M.E.	11	Glaser, R.	14	Ambron, S.	5
Langley, P.C.	6	Merrill, M.D.	5	Hooper, K.	7
Sobel, I.	7	Jonassen, D.H.	16		
Kaufman, R.	43	Reigeluth, C.M.	13	**5.0**	
Wright, P.C.	7			Anderson, J.G.	9
Swanson, R.A.	10	**4.0**		Jay, S.J.	9
Rothwell, W.	16	Landa, L.N.	5	Miller, P.L.	6
Zemke, R.	9	Gerlach, V.S.	15		
		Schmid, R.F.	7	**7.0**	
6.0		Simon, H.A.	12	Kirkpatrick, D.A.	7
Knowles, M.S.	5			Patton, M.Q.	7
Mager, R.F.	5				
* Campbell, D.T.	5	* Hartley, J.	5		

* **Isolates**

Title:

Mathematics Problem Solving within a Cooperative Learning Computer Structure

Author:

Bernard W. Arenz

Mathematics Problem Solving within a Cooperative Learning Computer Structure

Bernard. W. Arenz
Assistant Professor

Southwest Missouri State University
Elementary and Secondary Education
901 South National Avenue
Springfield, MO 65807

Paper Presented at the Annual Convention of the Association of Educational Communications and Technology, Orlando, FL, February, 1991.

Mathematics problem solving skills need to be developed through meaningful interactions. Computer simulations combined with cooperative learning methods provide students a context within which to construct an understanding of mathematics problem solving. This study examines the distribution and sequence of events that results in acquisition of those mathematics problem solving skills. Subject pairs of middle school age were observed solving mathematics problems presented by a computer simulated microworld. It appears that a complex set of strategies resulting in successful solution of the mathematics problems are developed by the subjects during the successful completion of the computer simulation and that those strategies may be acquired within pairs of subjects.

Mathematics problem solving continues to be a topic of concern for mathematics educators. Mathematical problem solving requires a student of mathematics to posses knowledge about a mathematical domain and to apply that knowledge with understanding to new and unique situations. Early attempts to teach mathematical problem solving through direct instruction met with mixed results (Gagne, 1965). Alternate views hold that the transition from novice to expert requires a person-to-person interaction in which the knowledge and understanding that comprises expert behavior is created and shared during meaningful problem solving activities (Bruner, 1960). The true test of mathematical understanding is in the application of mathematical thinking to solve appropriate problems (Schoenfeld, 1988). The Curriculum and Evaluation Stadards for School Mathematics (Standards) published by the National Council of Teachers of Mathematics in 1989 recommends that problem solving form the basis for mathematics instruction. The task of mathematics education, then appears to be one of providing students with opportunities to develop an understanding of mathematics problem solving that may be applied to a variety of problem solving situations.

Cooperative learning theory provides a methodology in keeping with the overall goal of developing mathematics problem solving skills consistent with the Standards. Cooperative groups provide a context that promotes and supports the person-to-person interaction needed for the development of problem solving knowledge and understanding. These groups, however, must engage in shared experiences and must be structured in such a way that students are encouraged to cooperate with each other in order to accomplish a common goal. This common goal develops positive interdependence between the students so that the success of each is dependent upon or enhanced by the knowledge and understanding of each other (Johnson, 1975). Thus cooperative learning provides the instructional framework for such a transmission of knowledge and understanding between learners.

The context within which that transmission takes place, however, must also be considered. Mathematics problem solving is thought to be best practiced within the context of familiar real-world problems. Students need to feel a connection to the problems that are presented. Perceived relevance of the problem by the problem solver is related to how closely the problem is grounded in the reality of the problem solver. The Standards (1989, p. 66) recommend that, "in developing the problem situations, teachers should emphasize the application of mathematics to real-world problems as well as to other settings relevant to the middle school students." In the case of mathematics text and workbooks the problems appear all to often to be contrived and not particularly relevant to the personal needs of the students. For the purposes of this study, a computer simulation that presents problems within the appropriate realm of the characters in a simulation is used to provide a relevant context for problem solving.

In summary, mathematics problem solving skills need to be developed through cooperative learning environment which allows individual problem solving skills to be reified and shared between students. The computer provides the students a context within which they are able to construct an understanding of mathematics problem solving. This study examines the process of learning mathematics problem solving skills, and the way in which a cooperative learning environment provides the forum for the transmission of those skills, all within a computer assisted learning context.

Purpose

The purpose of this study is to examine the process of learning mathematics problem solving skills and the way in which a cooperative learning environment provides the forum for the transmission of those skills between students within a computer assisted learning structure. The following three areas were investigated: (1) The distribution of mathematics problem solving characteristics used by students in solving mathematics problems, (2) the distribution of cooperative communications skills used by students in solving mathematics problems and (3) the acquisition of mathematics problem solving characteristics and cooperative communications skills.

The first question addresses the characteristics of mathematics problem solving. Students must proceed with a problem solving task with a complex set of actions that will lead them toward a solution of that problem. The actions are based upon the current resources, heuristics, control structures and beliefs of the student at the time of the activity (Schoenfeld, 1985). These mathematics problem solving characteristics develop over a period of time and may or may not directly relate to the computer based problem at hand. It is important to be able to categorize the problem solving strategies used by students in a computer based problem solving activity so that an accurate model of learner behavior can be built.

The second question is important to the examination of the effects that cooperative communications may have on the problem solving task. Difference in communication may take place when students are given the opportunity to work with each other on a problem solving task. In a traditional competitively structured classroom where all students compete with each other for grades, there may not be sufficient incentive for students to help each other. When new concepts are presented in a computer problem solving mathematics activity that are not easily understood, students may not feel comfortable asking peers for assistance and conversely, peers may be reluctant to give assistance. In the cooperative condition, however, the expectation is just the opposite (Johnson and Johnson, 1975). All students are required to help each other learn and to not do so would be a violation of class requirements and officially sanctioned behavior. What then happens when students are working together on a computer based problem solving task that requires them to communicate with each other in order to solve the problem?

The third question examines the acquisition of skills and understanding associated with the problem solving task. Each student brings a set of characteristics, as defined by Schoenfeld, to the problem solving task. Johnson and Johnson describe cooperative learning as a way in which skills may be shared with and learned from each other. It would be helpful to identify a pattern of problems solving characteristics and cooperative communication skills that become common to both participants during the computer based problem solving simulation.

The answers to these three questions should be helpful in dealing with the appropriate use of the currently scarce computer resources if learning mathematical problem solving skills can be shown to be enhanced by pairing students on computers while using cooperative learning strategies.

Significance

The way in which mathematics problem solving proceeds and how that may be taught within a classroom has long been a topic of interest to educators. The early hypothesis that problem solving across all activities was universally generalizable has given way recently to a belief that problem solving may be more content specific. This has in turn regenerated interest in mathematics problem solving as an activity requiring more finely tuned investigation. Mathematics problem solving activities that take place in a computer environment need to be explored within the broader context of mathematical problem solving but also as a special subset of that investigation. A concerted effort by the educational research community has not as yet been mounted in this area.

Rather sweeping changes in mathematics education are underway. Every student must "become a mathematics problem solver" and also "learn to communicate mathematics (Standards, p. 6)." The Standards are being used as an impetus to reexamine the mathematics curriculum in schools across the country and make changes in the way that mathematics instruction takes place in elementary and secondary schools. These changes are also have an affect on teacher training as colleges prepare preservice and inservice teachers for a changing mathematics curriculum. It is important to develop an increased understanding of the proper role that computers can play in assisting students to develop both mathematics problem solving and communication skills during this period of transition from a more traditional curriculum to one that mathematics educators move toward that is based upon the Standards.

The overall social environment within schools that technology threatens to fragment and isolate in the quest for higher achievement performance may be preserved through restructuring computer task structures. Cooperative learning techniques combined with judicious use of computers could prove to be a combination that provides the verifiable academic progress that is an expectation of schools at the same time as encouraging student to student interactions.

The increased interest to both give students computer access and also teach mathematics problem solving techniques has resulted in a number of computer programs being used for this combined purpose. With computers continuing to be a scarce resource, cooperative learning techniques may be an effective means of providing increased computer access at the same time as improved learning of problem solving skills.

The additional benefit may be that student-to-student social interaction goals may be advanced at the same time along with the other positive results of cooperative learning shown by other studies in this area. A demonstration of a cooperative computer learning environment improving performance would be valuable for educators primarily concerned with measurable performance. An environment that enables students to interact freely as they construct problem solving strategies may soften the use of technology for those critics of computers who are most concerned with the consequences of limited interpersonal interactions in computer assisted learning activities.

Literature

Problem solving activities in mathematics require skills and understandings that are often not readily apparent to the novice problem solver. It has been a goal of mathematics educators to provide students with the skills necessary for success in problem solving. Polya (1945) developed a set of problem solving steps that are claimed to be essential in problem solving and when properly applied to a mathematics problem: Step 1: understanding the problem, step 2: devising a plan, step 3: carrying out the plan; and step 4: looking back.

Schoenfeld extended Polya's work in order to develop a more complete description of the difference between a novice and expert mathematics problem solver. He studied the problem solving techniques of sophomore college students who had spent at least a year in the study of Euclidian geometry during high school. The subjects in this study were given typical problems in geometry and their attempts at solving the problems were videotaped. This investigation indicates that problem solving in this area may be characterized into four areas of: student knowledge and behavior; resources; heuristics; control and belief system.

In a cooperative learning situation these may be shared between the students so that the pooled resources may be applied to the task. The interactions of the students while engaged in cooperative computer problem solving exercises will be recorded.

The recorded interactions will then be examined for similarities and differences between this and problem solving interactions.

One area of investigation by Johnson, Johnson, Pierson and Lyons (1985) seems particularly intriguing, that of the use of controversy in goal setting. This research compared the use of controversy versus concurrence seeking as group goals. It was found that the groups with controversy as a goal showed higher achievement, motivation and a more accurate perspective than the groups with a goal of concurrence.

In the cooperative learning structure, the student and the instructor are brought together in a common bond based upon the humanistic needs inherent in teaching and learning. This bond is the manifestation of the love of learning expressed by Noddings. "The subject and student must be linked inseparably in the act of teaching. Linked to this passion for the material, but probably more important, is love of the acts of teaching and learning." (Noddings, 1984b, p. 157) This intrinsic connection between subject, teaching and learning continues as reflected in the relationships that both Noddings and the Johnsons propose as the basis for education.

"There is both an intellectual (or cognitive) representation and an affective one. If the activity is mainly intellectual, the intuition gives over its representation to reason, and objects are thought. If the activity is mainly moral, the intuitive representation is nearly pure feeling striking the motivational center with the force of its own being." (Noddings, 1984b, p. 73)

"I let the object act upon me, seize me, direct my fleeting thoughts as I scan the structures with which I may, in turn act upon the object. My decision to do this is mine; it requires an effort in preparation, but it also requires a letting go of my attempts to control. It involves a deliberate giving-over of subjectiveness." Noddings (1984b, p. 74) proposes that this giving over must be a conscious and deliberate act, "with a purposive quest for meaning and understanding."

Johnson & Johnson propose a curriculum that is guided by the need for cognitive, affective and content development needs of the learner. In this environment, each learner has the ability to grow and develop both socially and academically. This development is accomplished through activities that bind students together in cooperative groups. It is the group that forms the basis for both academic and social development. The distinguishing characteristic that separates the Johnsons' methodology from other group or team structures is the principle of positive interdependence. In positive interdependence, each student is responsible for the knowledge and welfare of the other group members. This is accomplished by setting external goal structures that require the common understanding of all group members. This could be contrasted with team work in which each member of the team is assigned a specialized task. Although the team needs each member to succeed in its primary function, team members are not responsible for acquiring the knowledge of other team members. In a learning environment, it is desirable that each team member is responsible for the knowledge and understanding of all other members. This will increase the helping behavior of all participants. Studies have shown that the gains in learning are highly correlated with explanation giving and receiving. Both the recipient and provider of information benefit from increased interaction that is focused on the knowledge and understanding necessary to complete the task at hand.

The emphasis on mutual responsibility for achievement sets up the cooperative relationship (Johnson, 1988). We all join together to realize the goals of education. In this way the teacher and students are dependent upon each other for their wellbeing. These elements combine to change both the motivation for learning and the way in which learning takes place. "The more cooperative a student's attitude, the more they see themselves as being intrinsically motivated." (Johnson, 1988, 3.12) The locus of control of the learning process is thus shifted from the external responsibility of the one-caring to the cared-for. The group process also provides an "...enduring effort aimed at achieving **satori**, or the sudden flash of enlightenment that comes after long, intensive, but successful effort." (3.15) These two elements combine to shift the focus of instruction from short term externally motivated skill acquisition to the pursuit of understanding and meaning.

Students must have the opportunity to work together so that they can reflect upon both the content and process used to acquire knowledge. In the competitive condition there is a negative incentive structure that limits or precludes the sharing of insight into the process. If the "secret" to success is shared, others will be more competitive thus lessening the chance for superior performance. An individualistic structure has unnatural avenues for the learner to share knowledge and understanding. Even with a strong internal tendency for the individual to assume a one-caring role the opportunity to fulfill that role may never present itself.

"Children and adolescents learn to view situations and problems from perspectives other than their own through their interaction with peers." The group work provides an opportunity for each member to take the one-caring role as well as the cared-for role. This shift allows students to look at themselves as others see them during the role shifts. (Johnson, 1988, 3.6)

Research in cooperative learning methods applied to computer assisted instruction activities indicate that computers may be used to further instructional goals that require group work. Students in the eighth grade were studied as they worked on a computer geography simulation in cooperative, individual and competitive modes of instruction (Johnson et.al., 1986). The researchers concluded from this investigation that there were positive effects for the subjects in the cooperative groups during a computer assisted simulated environment.

Cooperative learning research in a computer environment is limited, but sufficient enough to both indicate promise and yet require additional inquiry. Microcomputer interactions, however, are similar to interactions within other learning activities. Research carried out on the problem solving practice with microcomputers (Light et. al., 1987) found that 11 year old children paired together in a structured interaction were more likely to solve complex problems correctly. These results, it was pointed out, were similar to results obtained by Glachan & Light (1982) who conducted similar research using physical objects instead of the microcomputer.

It is reasonable to expect that although microcomputer environments do differ from more traditional learning environments, the interactions of students with each other in cooperative groups may be very similar. Current research into cooperative learning supports cooperative learning as a means for increasing academic achievement and a more positive attitude towards school and other students. This study is intended to extend the investigations into the relationship between mathematics problem solving, cooperative communications and a computer assisted instruction simulated microworld.

METHOD

This study incorporated group work, interaction characteristics, sequential analysis techniques and statistical results along with observational data. The interactions of pairs of students engaged in problem solving strategies and cooperative behaviors during the computer simulated mathematics problems were categorized and analyzed.

The instructional grouping in all cases was pairs of randomly selected students working together on the computer problem solving problem. Observations are broadly categorized as cooperative and problem solving. Sub categories for each of these characteristics are given in the section on observations. Analysis of the observation is divided into transitional probabilities and frequencies. Results are reported in the form of frequency graphs, tables and interaction matrices.

Treatment

Subjects worked on McGraw-Hill problem solving activities. This software creates an imaginary world in which the problems to solve are grounded. Subjects had the opportunity to solve problems that are both relevant to the characters in the story and lead to the solution of an overall problem. Preliminary investigation indicated that this simulated world is very motivating for the students. Problems in negotiating the simulated three dimensional space that the solution script occupies provided a varied experience that students found appealing.

Mathematics problem solving courseware published by McGraw-Hill and developed by Tom Snyder Productions was used as the basis of both learning activities. This series was built around a story line that had the student solving a mystery through a series of movements through a simulated building or geographic area. Each of the software packages were given grade level recommendations of 5 through 8. Each program started with a simple mathematics computation problem. The mathematics problems became more complex as the student(s) moved towards attainment of the simulation's overall goal.

Upon starting the program the subjects were faced with the task of using the arrow keys to effect physical movement throughout the simulated area. A series of statements were displayed on the screen that gave information to be used later in the solution of a question. This information may be stored in a "note book", a portion of computer memory that may be recalled at any time, by pressing the return key. When a question appeared the student must decide upon the appropriate method to arrive at the solution to the problem.

The first task was to select from the four mathematical operations of addition, subtraction, multiplication and division. After this selection was made the student had to find two appropriate number statements in the note book upon which the selected operation was performed. In the case of subtraction and division the order of the numeric statements was also critical to proper solution of the problem. The student can use the arrow keys on the computer to move back and forth through the list of previously stored statements. A key could be pressed to select a highlighted numeric statement into the solution set or the numbers could be typed directly on the keyboard. After entering all of the necessary numeric statements the computer performed the indicated mathematical computation. After examining the results of the computation, the student was given the choice of continuing computation or having the computer program check the solution for accuracy.

The later stages of the computer programs simulated journey required the student to use a multi step and sometimes iterative process to solve the problems that were presented. An additional complication was introduced with the inclusion of "red herrings", that is statements which appeared as plausibly useful but having no bearing upon the proper solution of the problems. These problems added the intermediate solutions to the note book as they were derived so that they could be used in later steps. The operations needed for solution were also varied during the intermediate stages.

There are four programs in this series of problem solving exercises. Each is targeted at a different student grade level. The story line was changed for each program so that the characters and situations provided novel and unique situations in which to solve problems. Each of the simulations was designed to give students experience with number strands and operations in various applications and presented within a variety of strategies.

A short introduction to the computer and overview of the assignment was given to all subjects before the initial session. This overview included directions about computer operation, procedures required by the software and general information about mathematics problem solving as related to the computer software. Instruction in general mathematics problem solving techniques or cooperative learning skills was not given.

The study was conducted over a five day period in which students worked on a mathematics tutorial program in problem solving. The problem solving computer program is self paced and combines some instruction with problems that need to be solved before moving to the next lesson. Students were told that this is an enrichment unit that is expected to help them in further math classes.

Each pair of subjects worked on a computer in a space that was part of the library but visually separated from other students by five foot high barriers. Each treatment session was a class period of about 45 to 50 minutes.

Subject Grouping: The subjects in this study worked in groups of two students. The group size of two was used because it is the minimum in which to allow the opportunity for sharing to take place. Subjects were students in 7th grade math classes in an upper Midwest urban school district. They were randomly assigned to the treatment from their regular classrooms. A total of five pairs were used as subjects for this investigation.

Interaction Characteristics: Each subject worked in pairs to establish group problem solving characteristics. Subject performance in the group setting was taped using video equipment for analysis of both verbal and physical interactions. Coding of the data was completed by two independent coders. Descriptions of interaction categories as well as tallies of interactions will be provided in the results.

Mathematics Problem Solving Activity: The computer simulation presented a series of mathematics word problems that were to be completed in order to progress through the simulation toward an overall puzzle or riddle. Each of the computational problems presented by the characters in the computer simulation were consistent with the general theme of the simulation and provided clues to the overall puzzle. The observation of the activities were coded into the broad categories of resources and heuristics and the further subdivided according to the follow subcategories:

Domain Knowledge
Facts, definitions, etc.
Algorithmic procedures
Routine procedures
Relevant competencies
Knowledge about rules

Local Heuristics
Review previous attempt
Request help message
Scroll through dialogue
Paraphrase computer message
Diagrams
Pencil and paper
Completed intermediate step
Completed problem
Select data for note pad

Cooperative Behavior

The subjects worked together in pairs on mathematics problems presented by the computer simulation. The simulation had an overall goal that became the focus for the pair so that both participants were mutually responsible for the achievement of the overall goal.

The basis for cooperation between individuals is their ability to communicate with each other in a manner that will allow each to contribute in a positive way toward a commonly defined goal. This communication is the transmission of both thoughts and feelings through words and actions that have a shared meaning. Each student must be able to send and receive accurately so that she may be able to "understand the other person's ideas, beliefs, feelings, opinions, reactions, needs, goals, interests, resources" and other things that may be related to the common objective (Johnson, 1975).

Cooperative Communications
Complete idea
Specific idea
Congruent verbal and nonverbal
Ask for feedback
Give feedback
Paraphrase partner message
Negotiate meaning or understanding

This set of cooperative learning communication skills formed the basis for interaction and observation during the treatment. Subject interactions were coded on the basis of these six categories of communications.

Sequential Analysis

The distribution of mathematics problem solving characteristics exhibited by the subjects while engaged in the mathematics problem solving computer simulation were determined by data recorded on the mathematics problem solving portion of the interaction matrix. The record of the number and duration of mathematics problem solving interactions exhibited by the subjects was analyzed for both frequency and duration. These are presented as a percent of the mathematics problem solving activities.

The distribution of cooperative learning skills exhibited by the subjects while engaged in the mathematics problem solving computer simulation was determined by data recorded on the cooperative learning skills portion of the interaction matrix. The record of the number and duration of mathematics problem solving interactions exhibited by the subjects were analyzed for both frequency and duration. These are presented as both a percent of the activities and percent of the time spent in cooperative learning skills.

Two independent coders tallied results. Inter-coder agreement is r = .83 using Cohen's Kappa test. Within characteristics differences were examined through comparison of percentages for the mathematics problem solving and cooperative communications characteristics.

The flow of mathematics problem solving characteristics and cooperative communications skills were explored through sequential analysis techniques (Bakeman, 1986). Each possible pair of observed interactions are evaluated for probability of occurrence and statistical significance. Probability maps were developed for the mathematical problem solving characteristics and cooperative learning communications. Statistical significance was tested using parametric techniques (z-scores) An interaction matrix was used to represent sequential relationships between pairs of events. Frequency tables of subject by time period were used to examine changes in paired events over time.

Descriptions of subject actions while engaged in the computer mathematics problem solving activity during the initial treatment and paired condition are reported. Instances that indicate a difference in subject behavior that indicate a change in either mathematics problem solving characteristics or cooperative communications are illustrated.

Both the statistical analysis and descriptive narrative are used as a basis for forming conclusions.

Observational Frequencies

A tabulation of mathematics problem solving characteristics and communications characteristics was used as the basis of further analysis. Transcriptions of the student-to-student dialogue and nonverbal interaction are provided. Examples of data recording sheets are included in the appendix.

Transitional Probabilities

The sequential relationship between the occurrence of events was analyzed through probability of occurrence of of those events. In the case of this study, the relationship of event A to event B is described by determining the probability event B following event A. This probability is then compared to the theoretical probability of this event sequence occurring at random. A statistical test was used to determine the statistical significance of the relationship between any two events.

Interaction Matrix

Interactions between subjects engaged in the mathematics problem solving computer activity was recorded on video tape during each session. These video tapes were then viewed by coders who have identified the time of onset and offset of each category of behavior given in the following observation matrix. Separate times are given for each cooperative learning communication skills and the mathematics problem solving skills. The cooperative communications skills and mathematics problem solving are anticipated to be nonexclusive behaviors so that the time of onset and offset of these overlapped each other. Coders were, however, required to differentiate between each of the behaviors exhibited by subjects within each of the cooperative and problem solving constructs.

Transitional Probability

The transitional probability schedule was derived from the actual data and is compared to the theoretical random probability of events. Each sequence of data has probability of occurrence based upon the actual observational data. This may be expressed as a calculated value between 0 and 1. Each sequence of data also has a theoretical probability based upon the likelihood of the events happening at random and has a calculated value between 0 and 1.

Each one tailed test with an $\alpha = .05$ is used for each possible sequence of transitional frequency in both the mathematics problem solving and cooperative learning strategies. A chart of transitional significance similar to the transitional frequency is used to display the statistical significance.

Observational vignettes

Narratives of interactions are used to provide insight into the nature of mathematics problem solving and cooperative learning strategies exhibited by the students. Mathematics problem solving within the context of the computer simulation is a complex activity. Descriptions of those interactions are essential to providing a full understanding of the problem solving activity. In addition, the character of cooperative interactions is as important as the frequency of interactions (Johnson, 1975). Vignettes that illustrate each of the problem solving and cooperative learning interactions exhibited by the subjects are included. These are used as a basis for conclusions as well as the statistical and frequency data.

RESULTS

These data were obtained by recording the interactions of five pairs of students. Each pair of subjects were randomly selected from a regular seventh grade mathematics class at an urban middle school in a medium sized upper Midwest metropolitan area. The subject pairs were assigned to a temporary laboratory setting within the instructional materials center of the middle school. The computer and video/audio taping equipment were set up each day before the arrival of the students and removed from the area after the students had completed each session. Each subject pair participated in five computer mathematics problem solving simulation sessions of about forty-five minutes in duration. Pencil and paper was provided for the subjects during each session

The flow of mathematics problem solving characteristics and cooperative learning skills were explored through sequential analysis techniques (Bakeman, 1986). Each possible pair of observational events were evaluated for probability of occurrence and statistical significance. Statistical significance is tested using parametric techniques (z-scores). An interaction matrix is used to graphically represent statistical significance of event pairs with only the event pairs with a $p < .05$ identified in table 1.

The likelihood of a pair of sequential events occurring in relation to all other event pairs was determined by using the Bakeman and Gottman statistical analysis technique previously described in the paper. These are given as onset, the first of two events, and the offset, the second of the events. Each onset may originate with either the first or second of the paired subjects with the offset the event undertaken by the other of the pair. Only the transitional events that occur between the subjects are tallied. Event sequences occurring within the same subject were not tallied during this analysis. The pair of events, number of occurrences during the sessions and the tally of occurrences by the first and second halves and first and second subject are listed.

An example of the dialogue between subjects for a sample of the event pairs is given in order to illustrate the way in which the subject pairs interacted while proceeding through the computer mathematics problem solving simulation. Descriptions of subject actions while engaged in the computer mathematics problem solving activity during the initial treatment and paired condition are also reported. Instances that indicate a difference in subject behavior that indicate a change in either mathematics problem solving characteristics or cooperative communications are identified. Interactions between the subjects and the media of instruction, the computer, will also be examined for possible influence on the outcome of the activity.

A tabulation of mathematics problem solving characteristics and communications characteristics is used as the basis of further analysis. Transcriptions of the student-to-student dialogue and nonverbal interaction that illustrate the significant transitional events are provided.

The following discussion is an example of the interaction between subject 1 and subject 2 for subject pair B. The transitional cooperative communications characteristic onset event "ask for feedback" and offset event "give feedback" were found to be significantly related. The frequency for the subject pair B over time is given in table 1.

Onset: Ask for Feedback
Offset: Give Feedback

Table 1. Frequency of transitional events for subject pair B.

		Subject	
		s1	s2
Period	1st	6	13
	2nd	4	4

In the first half of the mathematics problem solving simulation, subject 1 asked for feedback and subject 2 gave feedback six times. Subject 2 was responsible for the onset and subject 1 the offset of this event sequence 13 times. In the second half of the time spent by pair B, both subject 1 and subject 2 were responsible for 4 onset and offset actions. It appears that the actions of subject 1 and subject 2 were very dissimilar during the first half of their work with the mathematics problem solving simulation. The frequency of onset and offset of asking for and giving feedback became the same during the second half of the interactions.

In regard to asking and giving feedback, subject 2 initiated a question and subject 1 responded more than twice as often during the first half of their interactions. It appears that subject 1 had a greater understanding of the procedures necessary to move towards solution of the mathematics problems and was willing to share that understanding. During the second half of the interactions, both subject 2 as well as subject 1 increased their expertise in arriving at the solutions of the problem which resulted in the reduction of the incidents of information exchange between the pair of subjects.

The following sample event stream illustrated a typical interchange in the first and second halves of the computer activity. The transcription of dialogue between subject 1 and subject 2 is given in quotations.

1		s2	scrolls through the computer simulation
2		s2	selects information for the note pad
3	Give Feedback	s1	"Plus"
4	Give Feedback	s2	selects + 29 + 15 (=44)
5	Give Feedback	s1	"Not that far"
6	Give Feedback	s2	"That's right"
7	Give Feedback	s1	"Uhmm" (no)
8	Give Feedback	s2	"Yes it is"
9	Give Feedback	s1	"Uhmm; you're supposed to add this up" (points) "and subtract that" (points)
10	Give Feedback	s2	"Ah, ok, ya" selects + "go back up to 29+24"
11		s1	"There that's the answer"
12	Ask for Feedback	s2	"What about 15?"
13	Give Feedback	s1	"Doesn't matter....."
14	Ask for Feedback	s2	"And now this?" (refers to answer)
15	Give Feedback	s1	"No, no, no, now you subtract"
16	Ask for Feedback	s2	selects - "53"
17	Give Feedback	s1	"Ya go to 53"
18		s2	moves to 53

19		s1	"Up to 15"
20		s2	moves to 15
21	Give Feedback	s1	"Return" uses keyboard to move to answer
22		s2	selects answer
23	Give Feedback	s2	"Ya, got it right"

Second half of activity.

24		s1	scrolls through the character Bert
25		s1	selects information for the note pad
26	Give Feedback	s2	"Go to the note pad"
27		s1	"No problem, times that" points selects x 27 - x 6 = 16.....
28	Give Feedback	s1	"Now subtract"
29		s1	select 16.-2 +
30	Ask for Feedback	s1	"What did I do wrong?"
31	Give Feedback	s2	"I think it...." points to screen
32	Give Feedback	s1	"Oh, no times - subtract"
33		s1	subtract 46 - 16 = 30,000
34		s1	selects answer
35		s2	scrolls through the character
36		s2	selects information for the note pad
37		s1	selects information for the note pad
38		s2	selects information for the note pad
39	Ask for Feedback	s1	"What did you do so fast for?"
40	Give Feedback	s1	"Oh - add"
41		s2	chooses addition operation
42	Give Feedback	s1	"That plus that"
43		s2	selects the addends for computation
44	Give Feedback	s1	"Now divide that by 12"
45		s2	completes the division and then discusses the cursor location
46		s2	selects answer

These interchanges are examples of a shift of initiation in making sense of the current mathematics problem from one member of the subject pair to the other at two different times. In lines 3 through 9, subject 1 and subject 2 are engaging in a dialogue that does not require an explicit question in order to give feedback about the problem. Each in turn assists in making sense out of the mathematics problem and suggesting ways in which to proceed. The dialogue in lines 11 through 18 show the subjects engaging in an interchange in which subject 2 is asking subject 1 specific questions and subject 1 responds to those questions. In lines 26 through 34, during the second half of the coded events, subject 1 is requesting assistance from subject 2 and then receiving a response. The dialogue indicates that the individuals in subject pair B engaged in a cooperative communications about the current mathematics problem presented by the computer simulation in order to find the solution to the problem and that the responsibility for initiating inquires and providing responses may be dynamically adjusted during the problem solving episode.

Significant Event Interactions

Each of the subject pairs engaged in a number of event pairs that proved to be significant. Figure 1 gives a summary of the significant event pairs and graphically represents the categories into which those events fall. There are a total of eleven event pairs that are statistically significant for the subject pairs in this study. The following is a summary of those events.

Asking and giving feedback was a significant transitional event pair for subject pairs A, B and C. This indicates that part of the progress to solution to the mathematics problems may be the result of cooperative communications for these pairs but did not stand out for the other pairs.

Giving feedback along with the completion of the problem was a significant transitional event pair for the subjects pairs A, B and D. This may indicate that for these subjects, the giving of information is immediately linked to the solution of the mathematics problems.

The event of algorithmic procedure paired with another algorithmic appeared significant for subject pairs A, B and E. The determination of an appropriate algorithmic procedure to solve a problem seems to be an activity that requires interaction between both subjects in a pair to reach agreement.

An algorithmic procedure paired with a routine procedure was significant for subject pair E. All other subject pairs were able to move through the simulation without a significant number of algorithmic to routine procedures occurring. This may indicate a unique problem solving style demonstrated by the subject pair.

Routine procedure to giving feedback were significant transitional events for subject pairs B and D. For these subjects the execution of a routine procedure was followed by an unsolicited comment on the part of the other.

Scrolling through the computer simulation dialogue followed by selecting data for the note pad was a significant transitional event for subject pairs C, D and E. This appears to indicate that for these subjects the selecting of data after scrolling form an important sequence.

Paraphrasing the computer message with communicating a specific idea was a significant event pair for subject pair C, D and E. This may indicate that a characteristic of the problem solving style for this pair was to read back the computer dialogue and then articulate a idea that was related to the dialogue.

Completing a problem with scrolling through the computer dialogue was a significant event pair for subject pairs B, C, D and E. This may indicate that the subject pair A has a unique characteristic of the problem solving style that was not in evidence for any other pair making this an unimportant event for this pair. All other pairs demonstrated a significant engagement in the event pair.

Selecting data for the note pad followed by the articulation of a specific idea was a significant event pair for subject pairs A and D. It appears that for these subjects, after one of the pair entered data the other one of the pair would make a comment about that activity.

Selecting data for the note pad followed by scrolling through the computer dialogue was a significant transitional event for subject pairs C, D and E. For these two subject pairs, continued scrolling through the dialogue after entry of data formed an important characteristic of their problem solving strategy.

In summary, each of the subject pairs are somewhat unique in the way that they solve the mathematics problems presented by the computer simulation, however, they also show some common characteristics. Only the statistically significant event sequence of algorithmic followed by routine procedure was exhibited by only one subject pair. All other event pairs were commonly shared by one or more of the subject pairs.

Relative Distribution Of Events

Although the general pattern of problem solving seems to be fairly consistent, each of the subject pairs has developed different patterns of interactions in order to complete the mathematics problem solving activities in the computer simulation. Those patterns appear to be distributed in ways that reflect both their level of expertise in problem solving and their ability to work together toward common aims. Figure 2 illustrates both the similarities and differences in the way that the subject pairs approach the problem solving activity.

The general distribution for each characteristic category is similar for all subject pairs. This indicates that even though there are differences in the way that each pair of subjects goes about the task of solving the mathematics problems, there are common patterns of activity that lead toward that goal. It should also be noted that some of the traditional strategies for problem solving either were not used or made explicit during the activity.

General heuristics identified by Schoenfeld (1989) were almost entirely absent as part of the subject's problem solving repertoire while the domain knowledge of mathematics was very apparent. This may indicate that this computer software does not provide an environment in which these skills are important. Local heuristics that were a necessary component in the solution of the mathematics problems presented by the computer were both evident and appeared in greater number.

In summary, the relative distribution of mathematics problem solving activities indicates that subjects did not exhibit all of the problem solving characteristics under investigation in this study. Second, the overall distribution of events forms a consistent pattern for all of the subject pairs. Third, there are marked differences within the subject pairs in the relative distribution of event types when compared with other subjects.

The distribution in domain knowledge is fairly consistent in that all of the subject pairs seem to spend the most time with algorithmic procedures and routine procedures. Little time was devoted to the determination of facts or definitions. Instances of explicit action indicating relevant competence and knowledge about rules were relatively few.

Most of the domain knowledge category of activity is devoted to determination of the appropriate algorithmic procedures and routine procedures. There are differences in the proportion of algorithmic procedures and routine procedures between subject pairs. These differences are most apparent between subject pair A and subject pair C. Subject pair A engaged almost equally in determining algorithmic procedures and routine procedures. This ratio seems to indicate that for this subject pair it took no more time to determine the way in which a problem was to be solved, algorithmic procedures, as it was to carry out the solution of the problem, routine procedure. Subject pair

C, however, devoted a large proportion of their actions towards determining the appropriate way to solve the problem, algorithmic procedure, than to actually solving the problem, routine procedure. Other subject pairs in the study fell between these two extremes. They did, however, expend a greater effort in determining the appropriate algorithmic procedure than in demonstrating the routine procedure.

DISCUSSION

This study indicates the interactions of subjects who are engaged in the solution of mathematics problems presented to them through a computer assisted instruction simulation. Subjects in this study modified their interactions in ways that appear to ensure success in mathematics problem solving. Event sequences were modified between the first and second half of the activity that reflect the growth of understanding in both the local problem solving heuristics and the domain knowledge characteristics. Subjects in subject pair A at times would refer to the way in which the current problem was similar to an earlier problem. This notion of change and adaptation could also be seen in the way in which the cooperative communications event sequences were modified. Three of the pairs experienced a decrease in the total number and difference between the number of times that a sequence of questions and answers were necessary during their progress through a mathematics problem.

Subjects in the study maintained progress in solving the mathematics problems without direct instruction from the researcher or teacher. Some time was necessary in which the subjects struggled with the problems in order to construct an understanding of the way that the problem could be solved. The number of incidents in which it was necessary to construct understanding decreased between the first and second time periods in the study. Direct instruction was not necessary for successful completion of the simulation. This indicates that the subject pairs were able to construct their own knowledge and understanding of mathematics problem solving during the computer activity. This may be especially helpful for teachers who are would like to provide their students with mathematics problem solving activities that do not require constant monitoring or progress.

Subjects displayed differences in distribution of problem solving and cooperative communications characteristics. Although a great deal of similarity in the problem solving interaction events was present there was also a variation between the subject pairs. Some of the subjects spent more time on discussion of routine than algorithmic procedures than others. Some pairs devoted a larger amount of effort to master the intricacies of the computer software.

Individual differences existed within each of the pairs. Some members of the subject pairs contributed a greater number of events than others. This difference appears to be tied to the amount of knowledge of mathematics procedures and willingness to give information. The dominate individual, however, did assist the other subject in the pair develop strategies in problem solving that were reflected in a change in event distribution between the first and second halves of the coded events.

Consultation with each other was an accepted and often used strategy for determining appropriate responses to problems. Subjects were readily willing to engage in expressing their knowledge and understanding about the procedure for mathematics problem solving. These interchanges were in keeping with the model of cooperative communications developed by Johnson and Johnson. It appears that subjects who were not particularly adept at cooperative communications, as show by the number of coded events, became more like their partner in cooperative characteristics during the second half of the computer activity. This may indicate the the computer simulation could also have value in giving students the opportunity to learn and practice cooperative skills that could be used in other classroom activities.

Limitations

The sample size in this study makes general inferences about the distribution of problem solving characteristics problematic. Although the subjects were mixed race and gender, it would be presumptuous to assume that the observations made would necessarily reoccur in other students. Additional studies with a larger number of students combined with inferential analysis techniques would be helpful in drawing conclusions about the possible benefits for larger groups of classroom size.

No attempt has been made at a quantitative analysis of differences between subject pairs or within pairs. Differences in learning styles, learner control, ethnic or gender differences, attitudes and aptitude have not as yet been explored. There may be value in exploring these other areas to round out the picture of mathematics problem solving within the limitations of this computer simulation.

Summary

Mathematics instruction is in the processes of evolution towards a greater emphasis on problem solving with the Standards as one of the catalysts for that change. Cooperative learning methods combined with the appropriate use of computer technology may be able to play a part in this change by providing educators with another way in which to organize instruction and present problem solving challenges. This study indicates that for this limited case mathematics problem solving skills along with cooperative communications can be a positive supplement to more traditional instruction. Additional work in this area is indicated in order to assess the extent and magnitude of the learning taking place during the completion of the computer simulation.

References

Bakeman, R. & Gottman, J. M. (1986). Observing interaction: an introduction to sequential analysis. Cambridge: Cambridge University Press.

Balla, J., Gow, L. & Burton, P. (1986). Computers in the teaching process and software development: co-operation the key. Education and Computing, 2 (4), pp. 265-271.

Bruner, J. S. (1960). The process of education. Mass., Harvard University Press.

Collis, B. A. & Gore, M. (1987, February) The collaborative model for instructional software development. Educational Technology, pp. 40-44.

Cox, D. A. & Berger, C. F. (1985). The importance of group size in the use of problem-solving skills on a microcomputer. Journal Educational Computing Research, 1 (4), pp. 459-468.

Curriculum and Evaluation Stadards for School Mathematics (1989). Reston, National Council of Teachers of Mathematics.

Dansereau, D. F. (1988). Cooperative Learning Strategies. Learning and Study Strategies: Issues in assessment, instruction, and evaluation [Chapter] (pp. 103-120). Harcort Brace Jovanovich.

Fazio, R. P. & Berenty, F. J. (1983, October). Everybody wins in group computing. Science Teacher, 50(7), 56-58.

Gagne, R. M. (1965). The conditions of learning. New York, Holt, Rinehart & Winston.

Gallini, J. K. (1987). A comparison of the effects of logo and CAI learning environment on skills acquisition. Journal of Educational Computing Research, 3 (4), pp. 461-477.

Hartson, H. R. (Ed.). (1985). Advances in human-computer interaction. New Jersey, Ablex.

Humphreys, B., Johnson, R. T., & Johnson, D. W. (1982). Effects of cooperative, competitive, and individualistic learning on students' achievement in science class. Journal of Research in Science Teaching, 19(5), 351-356.

Hunt, N. P. & Allen, E. G. (1988, March/April). Fifth graders do "real work" on the computer. The Social Studies, pp. 78-80.

Jaques, D. (1984). Learning in groups. Dover, New Hampshire: Croom Helm.

Johnson, D. W. & Johnson, R. T. (1975). Learning together and alone: Cooperation, competition, and individualization. New Jersey: Prentice-Hall Inc.

Johnson, D. W. & Johnson, R. T. (1985). Cooperative learning: One key to computer assisted learning. The Computing Teacher.

Johnson, D. W. & Johnson, R. T. (1986). Computer-assisted cooperative learning. Educational Technology, pp. 12-18.

Johnson, D. W. & Johnson, R. T. (1986). Encouraging student/student interaction: Research matters to the science teacher. National Association for Research in Science Teaching.

Johnson, D. W. & Johnson, R. T. (1986a, January). Computer-Assisted Cooperative Learning. Educational Technology, 12-18.

Johnson, D. W. & Johnson, R. T. (1988). Cooperation in the Classroom. Minnesota: Interaction Book Company

Johnson, D. W., Johnson, R. T. & Stanne, M. (1984). Circles of learning. Washington, D.C.: ASCD.

Johnson, D. W., Johnson, R. T. & Stanne, M. (1986). Computer assisted instruction: A comparison of cooperative, competitive, and individualistic goal structures. American Education Research Journal.

Johnson, L. C. (1984). The effects of the "groups of four" program on student achievement. Paper presented to the Annual Meeting of the American Educational Research Association, 69th, Chicago, Il., March 31 - April 4.

Johnson, R. T., Johnson, D. W., & Stanne, M. B. (1986b). Comparison of computer assisted cooperative, competitive, and individualistic learning. American Educational Research Journal, 23(3), 384-392.

Kozma, R. B. (1987, November). The implications of cognitive psychology for computer-based learning tools. Educational Technology, pp. 20-25.

Lepper, M. R. (1985, January). Microcomputers in education: Motivational and social issues. American Psychologist, 40 (1), pp. 1-18.

Light, P., Foot, T., Colbourn, C. & McClelland, I. (1987). Collaborative interactions at the microcomputer keyboard. Educational Psychology, 7(1), 13-21.

Male, M., Johnson, D. W. & Johnson, R. T. (1986). Cooperative learning and computers: An activity guide for teachers. Educational Apple-cations, 194 Wedgewood, Los Gatos, CA.

Noddings, N. & Shore, P. J. (1984b). Awakening the inner eye: Intuition in education. New York: Teachers College Press.

Noddings, N. (1984a). Caring: A feminine approach to ethics & moral education. Berkeley and Los Angeles: University of California Press.

Noddings, N. (1986). Fidelity in teaching, teacher education, and research for teaching. Harvard Educational Review, 56(4), 496-510.

Okebukola, P. A. (1986). Cooperative learning and students' attitudes to laboratory work. School Science and Mathematics 86 (7), 582-590.

Polya, G. (1957). How to solve it: A new aspect of mathematical method. Princeton University Press, Princeton, New Jersey.

Schloss, P. J. & Schloss, C. N. (1986). Attributional consistency in CAI tutorials. Journal of Technology Systems, 14 (4), pp. 317-325.

Schoenfeld, Alan H. (1982, September). Problem perception and knowledge structure in expert and novice mathematical problem solvers. Journal of Experimental Psychology, 8 (5), pp. 484-494.

Schoenfeld, Alan H. (1985). Mathematics problem solving. Academic Press Inc., San Diego, Ca.

Schoenfeld, Alan H. (1988). Mathematics, technology, and higher order thinking. In R. S. Nickerson & P. P. Zodhaites (Eds.), Technology in education: Looking toward 2020. (pp. 67-96). New Jersey, Lawrence Erlbaum Associates.

Schuttenberg, E. M., McArdle, R. J. & Kaczala, C. M. (1985, April). Computer uses in schools: Research on what is and what should be. Educational Technology, pp. 19-22.

Shapiro, M. A. (1987, May). Control and interactivity in perceptions of computerized information systems. Paper presented to the Human Communication Technology Interest Group, International Communication Association, Montreal.

Solomon, C. (1986). Computer environments for children. Massachusetts, MIT Press.

Taylor, R. P. (Ed.). (1980). The computer in the school: tutor, tool, tutee. New York, Teachers College Press.

Vermette, S. M., Orr, R. R. & Hall, H. H. (1986, January). Attitudes of elementary school students and teachers toward computers in education. Education Technology, pp. 41-47.

Vockell, E. & Schwartz, E. (1988). The computer in the classroom. Santa Cruz, Mitchell.

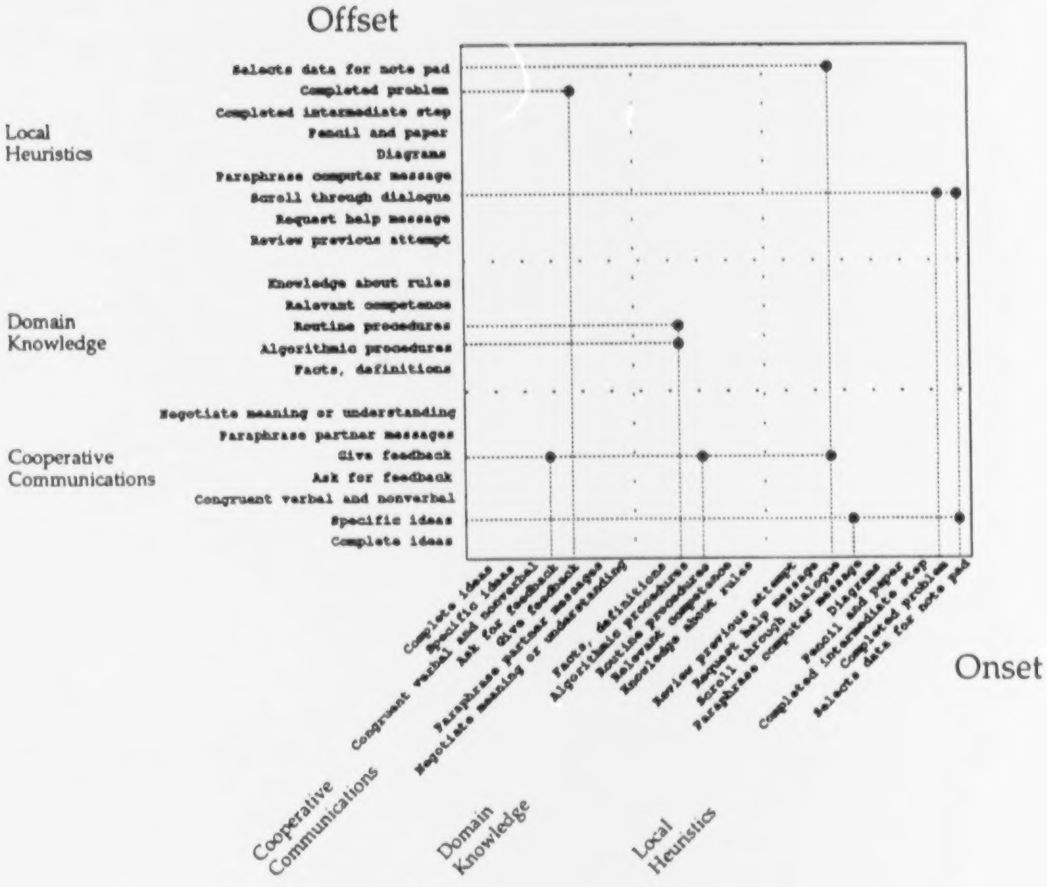


Figure 1. Transitional event onset to offset sequences for all subject pairs with $p < .05$ by event characteristic.

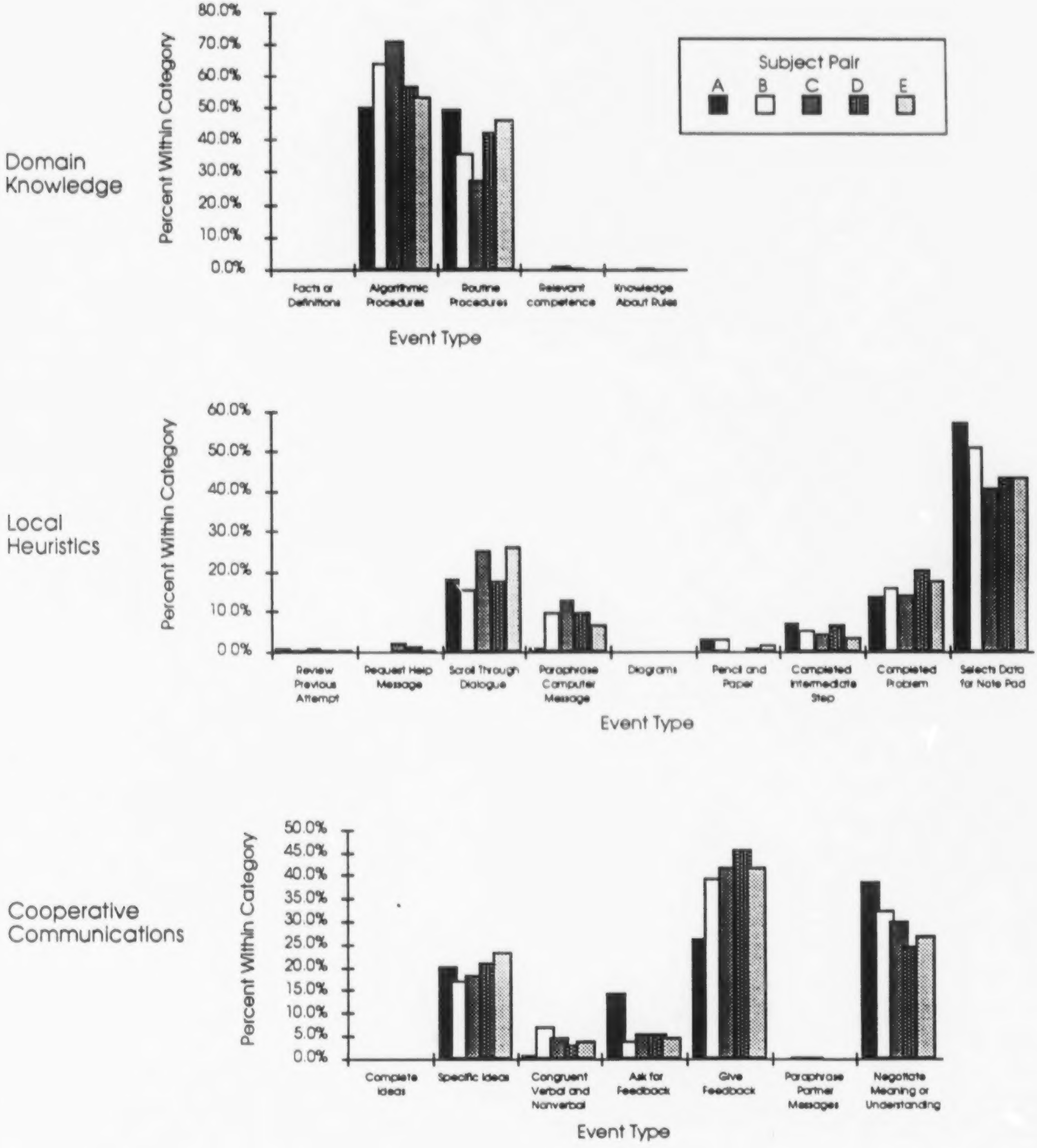


Figure 2. Relative distribution of events within mathematics problem solving and cooperative communications categories for all subject pairs in each event type.

BEST COPY AVAILABLE

Title:

Effects of Variations in Learner Control on Children's Curiosity and Learning from Interactive Video

Authors:

Marilyn Plavocos Arnone
Barbara L. Grabowski

ABSTRACT

With the emergence of interactive learning technologies, there are many questions which must be addressed concerning young learners. Such sophisticated technology combined with thoughtful instructional design has the potential for both encouraging achievement and stimulating important scholarly attributes such as curiosity and others aspects of motivation. The purpose of this study was to evaluate the effectiveness of variations in learner control (also referred to as lesson control) on children's level of curiosity and learning from computer-based interactive video (CBIV). The lesson content was art education and contained both facts and concepts. It was presented as a videodisc visit to the Everson Museum in Syracuse. A posttest only control group design was employed with 103 first and second grade subjects who were randomly assigned to one of three experimental conditions (designer control, learner control, or learner control with advisement) or a control group. The independent variable for this study was the degree of lesson control which the subjects had over the content. The dependent variables were the posttest scores in achievement and curiosity. Results indicated that children in the learner-control with advisement group scored significantly better in the achievement posttest than did the learner control subjects. They also tended to score higher on certain of the curiosity subscales.

INTRODUCTION

Increasing academic achievement while encouraging a young learner's motivation to learn is an important task for educators. More and more interactive learning materials such as computer-assisted instruction (CAI) and combuter-based interactive video (CBIV) are being developed for learners of all ages. The literature, however, is lacking in empirical research in interactive learning technologies on the effects of learner control (also referred to as lesson control) on children's achievement and motivation. This study used the following definition of learner-control: "Learner control can be described as the degree to which a learner can direct his or her own learning process . . . [The term] most often describes the instructional choices made during a particular lesson. By definition these choices can be made either by the instructional program (as originally defined by the designer) or by the learner during the presentation of materials (Milheim and Asbell, 1988, p.3)." The literature cited herein will include studies which examine degrees of learner control ranging from complete designer or program control (the learner makes no decisions about his/her learning process) to full learner control (the learner is in complete control of the available instructional options).

Most research involving interactive video and learner control have been conducted with sophisticated learners (Balson, Manning, Ebner & Brooks, 1985; Gay, 1986; Hannafin & Colamalo, 1987; Gay & Trumbell, 1988). Issues related to lesson control can have different implications for children (Hannafin, 1984). Although a completely unstructured format many prove challenging and acceptable for many adult learners, children with fewer prerequisite cognitive strategies may feel overwhelmed in an environment with so little guidance. Yet, overly restrictive conditions whether in a classroom setting or a designer controlled interactive video lesson, for example, leave little opportunity to stimulate the important scholarly attribute of curiosity. The literature in the two separate areas of learner control (as it relates to CBIV and CAI) and curiosity was helpful in designing a study which combines these factors.

A number of studies show that positive effects have been achieved with the use of interactive video (for example, Dalton, 1986; Bosco & Wagner, 1988; Hannafin M. & Colamalo, 1988; Abrams, A., 1986). Of particular interest are the studies in which lesson control was a factor. In the Hannafin study, subjects were randomly assigned to one of three treatments: linear control, designer control, or learner control. The designer control and learner control groups performed better than the linear control group. There were no significant differences, however, between the designer control and learner control groups. In the learner control treatment, students were advised up front of the recommended lesson sequence. This could be the reason why there were no significant differences between these two groups; essentially, all students in the learner control treatment followed the recommended path through the lesson. The potential for individualization and learning efficiency are cited as some of the benefits that accompany greater learner control (for example, Laurillard, 1984; Merrill, 1975). In CAI study involving 98 eighth graders, students performed better under learner control than program control; however, the form of learner control also offered the students a kind of advisement (Kinzie, Sullivan, & Berdel, 1988).

While some evidence exists that even secondary school-aged students can make thoughtful decisions in learner control situations (Robson, Steward & Whitfield, 1988), Carrier (1984) states plainly that there is little confirmation that learners make good choices when given the chance. One study which substantiates this claim showed that students selected fewer than the optimal number of examples of math rules to support their learning when in a learner control treatment (Ross & Rakow, 1981). The learner control group did not perform as well as the program control group which received examples based on need as indicated in a pretest. Time on task may also be a factor contributing to poorer performance as learner control students tend to spend less time in the lesson (Tennyson, 1980; Tennyson & Buttrey, 1980; Johansen & Tennyson, 1983). This would indicate that students choose to exit the lesson before they achieve mastery. Steinberg (1989) suggests that beginning level students in a particular subject area may lack the descrimination skills and subject-specific learning strategies necessary for learner control to be effective.

While studies such as the above have shown learner control to be associated with poorer performance, the addition of some form of **advisement** to a learner control lesson has been associated with positive results in increasing achievement (Tennyson, 1980; Johansen & Tennyson, 1983; Tennyson, 1984; Tennyson, Christensen and Park, 1984). Johansen & Tennyson (1983) used an advisement strategy which informed subjects of their performance relative to the mastery criterion, and suggested the amount of instruction needed. Subjects then could make their own decisions about the sequence and the amount of instruction based on their perceptual understanding of their needs. In the advisement learner control condition, subjects stayed on task longer than either the partial learner control group or the program control group and exceeded the mastery criterion level established for the lesson. This form of adaptive advisement was also found to be effective, appealing, and efficient in a study by Santiago & Okey (1990).

The appropriate level of learner control may be contingent on factors such as individual differences or prior conceptual understanding. In one study, subjects with low prior conceptual understanding of a content area made poor decisions relative to sequencing and learning strategies when presented with a high degree of learner control (Gay, 1986). Goetzfried & Hannafin (1985) also found that prior achievement affected performance.

With few exceptions, the studies cited above used college students or adult workers as subjects. Hannafin (1984), cited earlier, notes that older learners may be better able to utilize the learner control options because of their refined cognitive abilities. Younger subjects (elementary or junior high school age) may not yet have the necessary skills to make the most of such a lesson. Therefore, there is a potential danger in generalizing the results of the existing findings to the young learner. This exploratory study is an attempt to expand the base of empirical evidence in the area of learner control and children in a CBIV lesson.

Let us now turn to the subject of curiosity. There are those who would argue that exemplifying curiosity in learning is even more essential than focusing on specific subject areas. One author writes of curiosity, "Without it we are condemned to be ordinary. With it we have a shot at being a part of the future (Weintraub, 1986, p.160)." That children do well in situations in which they are allowed choices and encouraged to learn through active exploration are principles embraced by proponents of the Montessori method of teaching as well as by the National Association for the Education of Young Children (Calvert, 1986). Materials are developed which stimulate a child's interest and curiosity.

Curiosity is often associated with exploratory behavior. Berlyne (1960) identified two forms of exploratory behavior, specific exploration and diversive exploration. Diversive exploration occurs as a person seeks new experiences or relief from boredom whereas specific exploration is encountered in situations in which there is conceptual conflict often resulting in curiosity arousal. The curiosity motivates exploration which resolves the conflict. According to Berlyne, the arousal state that motivates the "quest for knowledge" and is relieved when the individual attains that knowledge is epistemic curiosity (1960, p.274). Curiosity is also described as an arousal state in which the individual desires to know more about self or environment (Maw and Maw, 1966). Cecil, Gray, Thornburg, and Ispa (1985) extend the definition of Maw and Maw and consider curiosity to be an arousal state that leads to and is a prerequisite for exploration, play, and creativity. In proposing a model of teaching and learning, another author defines curiosity as "the individual's desire to question or investigate (Parker and Engel, 1983)."

In one part of a study in which children were asked to look at pictures and respond by asking questions about the picture, it was found that high curious children asked more questions and had more independent ideas (multiple questions that represented the same idea were counted only once) than low curious children (Maw and Maw, 1964). Another study tested whether epistemic curiosity can be intensified by pre-questioning subjects and determined that there was a significant difference in mean scores between subjects who received questions and those that did not (Berlyne, 1960). Not only did questions heighten epistemic curiosity but they also served to facilitate retention of facts when subsequently subjects encountered the questions associated with the facts.

Stimulating curiosity is an important responsibility of both parents and educators. The hypothesis that high curious children show better overall social adjustment than low curious children was accepted in one study of 577 fifth graders from New Castle County, Delaware (Maw and Maw, 1965). Other findings of the same study included that high curiosity children tend to be more consistent in thought processes as well as more creative and flexible, and high curiosity children seem to be more self-sufficient than low curiosity children.

To operationalize the definition of curiosity, the authors of this study used Maw and Maw's definition as it relates to school children. Maw and Maw went to great lengths to define the construct including informal and formal inquiries, review of the literature, and an examination of the historical development of the term. It seems to incorporate much of what was cited earlier, especially the work of Berlyne. "...curiosity is demonstrated by an elementary school child when he:

1. reacts positively to new, strange, incongruous, or mysterious elements in his environment by moving toward them, by exploring, or by manipulating them.
2. exhibits a need or a desire to know more about himself and/or his environment.
3. scans his surroundings seeking new experiences.
4. persists in examining and exploring stimuli in order to know more about them (p.31)."

This study focused on #1, #2, and #4 above. Number 3 is associated more with diversive curiosity whereas this study was more concerned with epistemic curiosity.

Learner control and curiosity is a challenging issue. For example, if the child is given a learner controlled lesson in order to facilitate curiosity and exploration, might he/she flounder for the lack of direction? Given a totally designer controlled lesson, might the child's potential to be curious about learning be stifled? Can interactive learning technologies such as CBIV be utilized to both stimulate curiosity and provide adequate direction and guidance? The purpose of this study is to investigate the comparative effectiveness of variations in learner control on children's curiosity and learning about art.

The null hypotheses for this experiment are (note that the hypotheses relate to content-specific achievement and curiosity):

1.) There will be no significant difference in achievment scores demonstrated in a posttest between control group subjects and subjects in any of the three experimental groups.
2.) There will be no significant difference in curiosity demonstrated in a posttest between control group subjects and subjects in any of the three experimental groups.
3.) There will be no significant difference in the achievement level of students receiving a designer control lesson and comparable students receiving a completely learner control or learner control with advisement lesson.
4.) There will be no significant difference in the achievement of students receiving the completely learner control lesson and comparable students receiving a learner control with advisement lesson.
5.) There will be no significant difference in curiosity of students receiving a designer control lesson and comparable students receiving the same lesson in a completely learner control or learner control with advisement condition.
6.) There will be no significant difference in curiosity of students receiving a completely learner control lesson and comparable students receiving a learner control with advisement lesson.
7.) There will be no significant differences found in the correlations between achievement and any of the curiosity measures.

METHODS

Subjects

The sample consisted of 103 first and second grade students who attended a moderately sized public elementary school in upstate New York. Only individuals who were given parental permission participated in the study. No attempt was made to group subjects by ability or reading level. Two subjects were dropped from the study; one had severe mental and physical handicaps which made it impossible for the subject to respond to the touch screen, and the other did not participated due to illness.

Description of Lesson

An interactive video lesson was developed using a Sony Laservision videodisc player interfaced with a MS-DOS compatible computer, a touch-sensitive screen, and headsets. The lesson was programmed using ICON Author. A combination of motion video, slides, and computer graphics was used in the presentation. A subject matter expert in art education provided the content for the treatments. The lesson was designed as a visit to an art museum. The subject matter expect had previously served as Curator of Education for the museum and it was possible to acquire many visuals for the treatments. Prior to the lesson, each subject had the opportunity to practice using the touch screen. The lesson itself included a general introduction and three segments on ceramics, sculptures, and paintings. The content involved both facts and concepts and the lesson provided opportunities for practice, feedback, and remediation. The subject matter expert reviewed the posttest for content validity and determined that it adequately represented the content presented in the lesson. Another expert from Syracuse University Division for the Study of Teaching who works with talented and gifted children reviewed the posttest for evidence that it contained measures which addressed the aspects of curiosity as defined by the study.

The lessons were designed for subjects with limited reading experience and contained no text. A narrator was used where text would have been necessary. The use of a touch screen in place of a keyboard further simplified the young learners' task of responding.

CBIV Instructional Treatments

Three treatments were developed for this study: designer control, learner control, and learner control with advisement. All treatments contained the same essential content and all provided opportunities for practice, remediation, and feedback. Two instructional designers provided judgements on the appropriateness of the treatments in reflecting the three constructs as described below.

> *Designer Control*: Subjects receiving the designer controlled lesson followed a linear path through the lesson. These students received practice items, feedback, and were automatically branched to a remediation segment after a second incorrect response. Following remediation, the practice question was posed again; if the student still responded incorrectly

he/she was given the correct response and moved to the next item.

Learner Control: In the learner control group, subjects were given the opportunity to sequence the material in any way they preferred. Whether to review segments where practice items had been missed was also a decision left to these subjects. Students had the opportunity to omit entire sections or sub-sections, or opt out of the lesson altogether, if they so desired. Students controlled their own pace through the lesson. Additionally, subjects in this group could freeze images on the screen. We called this a "STOP and LOOK" routine since whenever a particular icon was present, the subject could freeze the image on the screen to explore it more closely. Subjects were familiar with the icon since they had been exposed to the icon and its meaning in the pre-lesson practice session.

Learner Control with Advisement: Subjects in the learner control with advisement group received the same opportunities to explore the lesson as the completely learner control group. However, certain "advisement" strategies were also employed in this group which provided some guidance while at the same time encouraged curiosity. For example, if a student decided to skip out of a section, he/she would receive this advisement: "Are you sure you want to end the lesson? This next section is very interesting. You might really enjoy it" or "Aren't you going to wonder about what you'll be missing?" Care was taken not to instill fear of evaluation as a motivation to continue the lesson since it would not be conducive to stimulating curiosity. While subjects in the learner control group could take advantage of the "STOP and LOOK" routine whenever they saw the associated icon appear on the screen, subjects in the learner control with advisement group were advised by the narrator to take advantage of the "STOP and LOOK" routine to explore the images more closely. The final form of advisement was in the form of a "STOP & THINK" routine which generally was preceeded by a question to arouse curiosity. For example, after presenting some interesting information about a painting that generally intrigues young children, this question was posed: "Do you wonder how you can tell this from looking at the painting? STOP and THINK about it! Then, touch the screen when you are ready to find out." All audio and visuals were programmed to freeze at that point and would resume only when the subject touched the screen to proceed. Since the learner had no control over the freezing of the screen, this was considered an advisement strategy and not a learner control option.

Other than the advisement, the instructional content remained the same as all other groups.

Research Design and Data Analysis

To test the hypotheses, this study employed a posttest only control group design. Subjects were randomly assigned to one of three experimental conditions in which they received a lesson followed by a posttest, or to a control group in which they received only the posttest. The research design is diagramed in Figure 1.

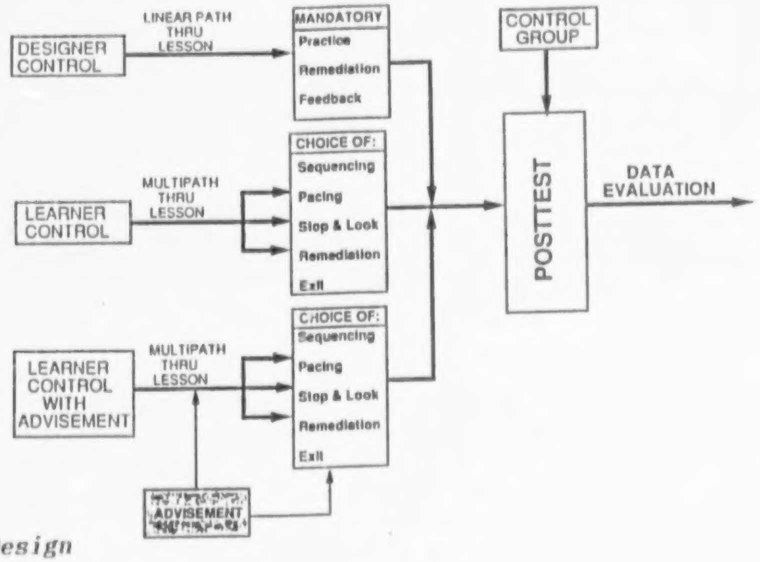

DESIGNER CONTROL
LINEAR PATH THRU LESSON
MANDATORY
Practice
Remediation
Feedback
CONTROL GROUP
LEARNER CONTROL
MULTIPATH THRU LESSON
CHOICE OF:
Sequencing
Pacing
Stop & Look
Remediation
Exit
POSTTEST
DATA EVALUATION
LEARNER CONTROL WITH ADVISEMENT
MULTIPATH THRU LESSON
CHOICE OF:
Sequencing
Pacing
Stop & Look
Remediation
Exit
ADVISEMENT

FIGURE 1
Research Design

Random assignment and the control group were used to control for differences in prior knowledge of museums and ability level. One-way analysis of variance was used to determine differences among the four groups on achievement (Hypotheses #1, #3 and #4) and again for curiosity (#2, #5 and #6). A .05 probability level was selected to determine significance. Since this was an exploratory study, the first in a series, a least conservative measure, Tukey's Least Significant Differences (LSD) test, was selected to analyze the differences between means in the follow-up tests.

Finally, correlations were run to test Hypothesis #7 using Pearson Product Moment Correlation Coefficients to determine the relationship between the three curiosity measures and achievement.

Procedure

Subjects were administered the treatments individually. In a pre-lesson exercise, a treatment administrator familiarized the subject with how to use the touch screen and introduced several icons they would encounter during the lessons.

Posttest responses were recorded both on audio tape and noted on a paper and pencil instrument. In one sub-section, the computer was used to track responses.

A number of provisions was made to protect the study's internal validity. Subjects were blind to which treatment they were receiving and were not informed about the nature of the experiment. So that all subjects would have the sense that they had received the same experience, all subjects including the control group participated in the same non-academic interactive video activity following the posttest. Since all the treatments were somewhat novel (i.e. all were interactive) and all ended with the same "fun" activity, the possibility for the Hawthorne effect to occur was reduced. To control for experimenter bias the authors of this study did not administer the treatments. Three administrators were selected to introduce the children to the lesson and to interview them for the posttest. The administrators were assigned treatments on a rotation basis so no one was responsible for just one treatment. All administrators participated in a training session on treatment and posttest procedures prior to the experiment and a beta test was conducted to assess the inter-rater reliability of the test. Raters came to 100% agreement on the rating by the end of the training session. There were three complete interactive video stations set up at the school which allowed for three students to participate at one time. The lessons and the posttest were reviewed by experts for evidence that the constructs defined in the study were adequately represented in the lessons and instruments.

Measurement

Part of the challenge of designing measures to test the effectiveness of the treatments had to do with the type of content that was to be presented. "Pat" answers would be inapproriate in an area such as art education and museums. In measuring achievement, we needed to devise an instrument that gave the child the opportunity to be more expressive while demonstrating that he/she had indeed acquired the new information. In measuring curiosity towards the content, an instrument that gave the child more freedom to explore possible answers was needed. Maw and Maw (1964) developed a number of tests to measure curiosity, one of which included presenting pictures to students. The student was be given one statement about the picture and then asked, "What else would you like to know about this?" Although Maw and Maw's first preliminary study showed that fewer questions were asked by students when too many pictures were shown together, their second study was modified to include fewer pictures. In that study, it was found that the mean number of questions asked and the mean number of independent ideas expressed in the questions was greater for the high curious children than for low curious children. This type of test seemed appropriate especially in the content area of art where presenting images would seem natural.

Achievement: The first section of the posttest, then, consisted of 8 items which measured achievement. To control for order effect and fatigue, the computer was programmed to randomly generate the test items for each subject. Because of the open-ended nature of the questions, we found that it was necessary to distinguish between responses which were specific to the lesson (lesson-related) and those which were not (lesson-unrelated). The control group had the opportunity to list numerous responses which had no bearing on the actual instruction (for example, in an item which represented a concept such as portrait, the control group child could increase number of responses merely by stating the various objects in the picture). To a lesser extent, subjects in other groups also provided such unrelated responses in addition to lesson-related responses. In the achievement measure, the subject was presented video still images of aspects of the museum encountered in the lesson such as paintings, sculptures, textures, and ceramics. For each item, the administrator requested the subject to tell him/her everything the subject knew about the picture. Responses were recorded on a paper and pencil instrument and also audiotaped. After responding, the subject touched the screen to proceed to the next image. The rationale was that prior to the treatments, the subjects would demonstrate little knowledge of the material and have limited vocabulary. Following the treatment, the subjects should be capable of demonstrating more and richer responses. For example, in the control group, a student might refer to a painting of Niagara Falls as "some waterfalls" whereas post-treatment response in the experimental condition might be "That's a landscape painting. It shows the natural scenery." The instrument was devised to take somewhat qualitative data and give it points based on 1) increased use of vocabulary and understanding of concepts introduced in the treatment 2) increase in the number of responses 3) recall of specific information.

Curiosity: Three sub-scales measured curiosity toward the content area. Subscale A was curiosity towards museums, Subscale B represented the number of questions the child generated when presented with a previously unencountered image, and Subscale C measured the time a child persisted in exploring new stimuli (indicated by the length of time the subject examined a new image before moving on to the next). Subscales B and C measured aspects of curiosity as defined in parts 1, 2, and 4 of the construct. As in the achievement section of the posttest, the computer was programmed to randomly generate the test items for subscales B and C. It should be noted that the number of items in the subscales was not large. This was a decision based on Maw and Maw's discovery that the more items asked the fewer the number of questions which were elicited from the learner. In Subscale B, the subject was rated on the number of questions they asked the administrator and on the number of independent ideas expressed by those questions. In other words, the more questions and more ideas--the more curious the subject had become towards the content area. Although the images presented in the posttest were art-related, subjects had never encountered the specific images in the treatment. The increase in questions, therefore, should not be related to the increase in information they had acquired, but rather an increase in curiosity brought about the treatment of the content.

Another measure of curiosity for this study which pertains to exploring new stimuli was the length of time a student stopped the videodisc in the STOP and LOOK mode during the lesson itself. This information was tracked by the computer. Since students in the designer control treatment did not experience this control option, there were five STOP and LOOK items on the posttest for all groups. These items comprised Subscale C.

The computer was also used to track each student's progress through the lesson. The authors plan to use this information to assess which learner control attributes of the interactive video technology the students' utilized most to explore the content, and which treatment was most conducive to encouraging curiosity about the content. Through analysis of the child's path through the lesson, we will also determine how thoroughly the child explored his path through the lesson (which relates to both #1 and #4 in the operationalized definition).

RESULTS

Achievement: Differences between the treatment groups on achievement were analyzed on two variables, lesson-related responses and lesson-unrelated responses using one-way ANOVA. Descriptive results for the lesson-related responses are reported in Table 1. Means ranged from a low 5.00 for the control group to a high 16.08 for the learner-control with advisement group. Differences between the means for each treatment group on the two variables are shown in Figures 2-4.

TABLE 1
Means and Standard Deviations for Lesson-Related Achievement Scores By Treatment Group

Treatment Group	N	M	SD	RANGE
Learner Control with Advisement	25	16.08	6.54	4-35
Learner Control	27	13.35	4.49	3-23
Designer Control	25	13.89	4.98	4-22
Control	24	5.00	2.38	1-9

TABLE 2
Results on ANOVA by Lesson-Related Responses

Source	df	SS	MS	F	p
Treatment	3	1725.10	575.03	24.49	.0001

Results from the ANOVA indicated a difference between treatments at p<.0001 (See Table 2). Follow-up tests showed that differences were found, as expected, between the control group and all three experimental groups. Of interest to this study was the difference between the learner-control with advisement group scoring significantly higher (+2.731 p<.05) than the learner control group. No significant differences were found between the designer-control group and the learner-control or the learner- control with advisement groups. Learner-control with advisement lessons, therefore, resulted in the greatest amount of achievement followed by designer-control and learner-control, and the control group.

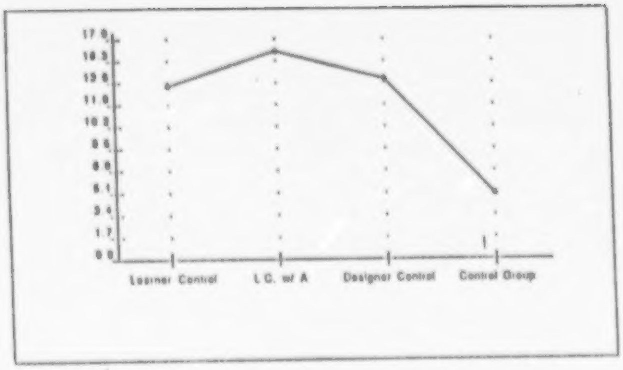


FIGURE 2
Mean Scores on Lesson-Related
Achievement by Treatment Group

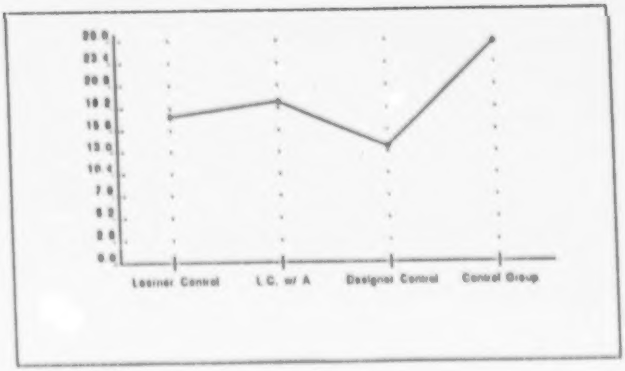


FIGURE 3
Mean Scores on Lesson-Unrelated
Achievement by Treatment Group

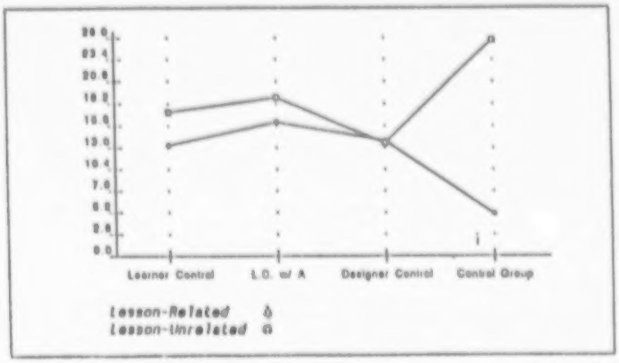


FIGURE 4
Mean Score Achievement Comparisons

Descriptive results for lesson-unrelated responses are reported in Table 3. Differences between the means for each group are shown in Figure 3. Means for unrelated responses ranged from a low of 13.55 for designer-control and a high of 25.76 for the control group. The spread of these responses was quite different than for the lesson-related responses. This spread is accounted for by the control group. Since the control group did not have the lesson, it is reasonable that the number of irrelevant (unrelated) responses would be greater than relevant (lesson-related) ones. Except for the expected tremendous difference in the control group, the analysis of the lesson-unrelated achievement scores yielded similar trends as the lesson-related scores as can be seen in Figure 4.

Results from the ANOVA indicate a difference between the treatments at p<.0034 (See Table 4). Follow-up tests showed that differences were found only between the control group and all other experimental treatments. No other significant differences were found.

BEST COPY AVAILABLE

TABLE 3
Means and Standard Deviations for Achievement Scores by Lesson Unrelated Responses by Treatment Group

Treatment Group	N	M	SD	Range
Learner Control with Advisement	25	18.96	12.79	5-53
Learner Control	27	17.16	9.84	3-45
Designer Control	25	13.55	9.25	2-39
Control	24	25.76	13.68	10-70

TABLE 4
Lesson-Unrelated ANOVA Results

Source	df	SS	MS	F	p
Treatment	3	1922.35	640.78	4.85	.0034*

* Indicates significance at <.05

Curiosity: Analysis of the difference between the groups on curiosity was conducted on each subscale, Subscale A (curiosity towards museums), Subscale B (number of questions generated) and Subscale C (persistence in exploring new stimuli). Subscale B was also further divided into total number of questions generated (Subscale B1) and number of independent ideas represented by those questions (Subscale B2). Descriptive results for all curiosity subscales are reported in Table 6.

TABLE 5
Means and Standard Deviations for Curiosity Measure Subscales by Treatment Groups

	Subscale A			Subscale B1			Subscale B2			Subscale C		
Trt	N	M	SD	N	M	SD	N	M	SD	N	M	SD
LC	26	3.42	1.77	27	13.38	7.17	27	12.89	6.81	24	57.58	37.20
LC/A	23	3.74	2.32	25	16.76	6.39	25	15.56	6.06	23	55.57	29.25
DC	22	3.18	1.33	25	14.60	9.48	25	13.96	9.22	23	52.70	23.21
C	20	2.60	1.14	24	15.90	9.24	24	14.93	7.98	22	75.97	45.02
TOT.	91	3.26	1.74	101	15.12	8.13	101	14.30	7.50	92	60.25	35.16

Results from the ANOVA for each subscale are reported in Table 6. No significant differences were found in curiosity as measured by curiosity towards museums (Subscale A), total number of questions generated (Subscale B1), number of independent ideas represented by the questions (Subscale B2), or persistence in exploring new stimuli related to art (Subscale C).

TABLE 6
Results of ANOVA by Subscale

Source	df	SS	MS	F	p.
Subscale A	3	14.82	4.94	1.66	.1816
Subscale B1	3	170.22	56.74	.85	.4676
Subscale B2	3	105.78	35.26	.61	.6091
Subscale C	3	7413.94	2471.3	2.07	.1100

Follow-up tests showed significant differences only on Subscale A between learner-control with advisement and the control group, and for Subscale C (persistence) between the designer control and the control group. Further examination yielded an interesting although an only partially significant trend on the last difference. The mean time persisted in attending to new stimuli decreased as the amount of instructional control increased. The most curious students on this measure were the control group students who had not even been exposed to the lesson treatment. The least curious were subjects in the designer control lesson who had the greatest amount of instructional control imposed on them.

Achievement and Curiosity: Lesson-related, lesson-unrelated, and total achievement scores were correlated with each curiosity subscale measure and are reported in Table 7.

TABLE 7
Correlation Coefficients Among Achievement and Curiosity Measures

	CURIOSITY			
ACHIEVEMENT	A	B1	B2	C
Lesson-Related	.29*	.26*	.22*	.02
Lesson-Unrelated	.32*	.47*	.52*	.11
Achievement Total	.47*	.59*	.62*	.12

(* indicates significance at p<.001)

The relationship between Achievement and Curiosity measuresB1 (total number of questions generated) and B2 (number of independent ideas represented by the questions) were the only positive correlations.

DISCUSSION

This study proved more useful in examining links between achievement and learner control among very young children than it was in learning more about the relationship between learner control and curiosity relative to specific content. Achievement in terms of recall was of course much easier to measure in this study. Examining curiosity for the content, on the other hand, proved much more difficult and the results in the area of persistence almost seemed counter-intuitive.

In lesson-related achievement, the learner-control with advisement lesson resulted in the greatest amount of achievement followed by designer-control and learner-control. Naturally, the control group, having received the posttest only, received the very lowest scores. What this seems to indicate is that some structure to a lesson in the form of advisement was better than allowing this age group free reign in discovering the content. Learner-control with advisement also provided the most opportunity for interaction which has been shown in previous studies to result in greater learning of factual information (Schaffer and Hannafin, 1986). It was expected that the learner-control group with no guidance whatsoever would have the lowest achievement scores, but in actuality, this group's scores were only marginally less than those of the designer control group. Whether what was lost in the structured presentation of content provided by the designer-control lesson was made up for in the learner-control lesson via opportunities to explore at will thereby increasing interest and recall of the information remains unanswered at this point. A thorough analysis of the students' paths may provide further clues on this point.

The young learners who took the learner-control with advisement lesson also generated the most questions and the most independent ideas in those curiosity subscales followed by learners taking designer-control and learner control lessons respectively. The control group settled somewhere in the middle reflecting an average curiosity level. There are several possibilities for the overall lack of significant results in these curiosity subscales. Perhaps, the construct was not adequately captured in the measures which were designed specifically for this study (although they were adapted from previous studies). It could also be that additional strategies which would encourage curiosity need to be incorporated into the lessons themselves, and that the differences between treatments with respect to those strategies need to be strengthened in order to increase the chances of sensing an effect.

Curiously (no pun intended), there was one interesting significant difference related to curiosity between the control group and all other groups. The finding involved the persistence subscale and initially seemed counter-intuitive. It indicated that students who were not exposed to the art education content were more apt to attend longer to stimuli that was totally new to them (that is, although all students were attending to new stimuli at this point in the posttest, the other groups had at least been exposed to the subject matter area in the treatment and thus the type of stimuli was at least familiar). Naturally, one might argue a rival hypothesis for this result being attributable to fatigue or even boredom. The students in the control group were given the posttest only while the others had been through a lesson as well and may well have been ready to move on by this time. The trends indicated by these scores among the four groups, however, may be worth further investigation for if, in fact, learners persist less in attending to new stimuli as instructional control increases, this would suggest potential instructional strategies for young learners. The generalizability of the study is of course limited in that the results were specific to the subject matter domain.

Possibly, a future study using the same interactive video materials, with modifications, should be conducted in which existing validated reliable measures of general curiosity are taken. A study of this sort would provide the opportunity to investigate the differences between high curious and low curious children regarding their preference for and achievement in various learner control conditions. The findings of such a study could provide information to designers of CAI and CBIV to tailor instruction and opportunities for learner control to the intrinsic needs of the students. For example, a low-curious student may hypothetically perform better in a more structured format such as designer-control but certain strategies might then be embedded in that treatment specifically to encourage curiosity in the low curious child.

Is there an optimal balance between amount of instructional control and achievement and motivation (eg. curiosity)? What other factors may be involved such as cognitive styles, general curiosity, preferences, etc.? What can educators and designers of interactive learning technologies do to encourage children's quest for knowledge and intrinsic motivation to learn? The few significances that were found in this exploratory study combined with the trends in means indicates to us that more research with young learners is warranted in the area of learner control and achievement and motivation.

REFERENCES

Abrams, A. (1986, January). Effectiveness of interactive video in teaching basic photography skills. Paper presented at the Annual Convention of the Association for Educational Communications and Technology, 14 p.

Balson P., Manning D., Ebner D., & Brooks, F. (1985). Instructor-controlled versus student-controlled training in a videodisc-based paramedical program. Journal of Educational Technology Systems, 13(2), 123-130.

Berlyne, D. E. (1960). Conflict, Arousal, and Curiosity. New New York: McGraw-HIll Book Company.

Bosco, J. & Wagner, J. (1988). A comparison of the effectiveness of interactive laser disc and classroom video tape for safety instruction of General Motors workers. Educational Technology. 28(6), 15-22.

Calvert, P. (1986, November). Responses to guidelines for developmentally appropriate practice for young children and Montessori. Paper presented at the Annual Meeting of Young Children, Washington D.C., 12 p.

Carrier, C. (1984). Do learners make good choices? Instructional Innovator, 29(2), 15-18.

Cecil, L., Gray, M., Thornburg K., & Ispa J. (1985). Curiosity-exploration-play-creativity: The early childhood mosaic. Early Childhood Development and Care, 19, 199-217.

Dalton, D. W. (1986). The efficacy of computer-assisted video instruction on rule learning and attitudes. Journal of Computer-Based Instruction, 13, 122-125.

Gay, G. (1986). Interaction of learner-control and prior understanding in computer-assisted video instruction. Journal of Educational Psychology. 78(3), 225-227.

Gay, G., Trumbell, D., & Smith J. (1988). Perceptions of control and use of control options in computer-assisted video instruction. Tech Trends, 33, 31-33.

Hannafin, M. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction. Journal of Instructional Development, 7(3), 6-9.

Hannafin, M. & Colamalo M. (1987). The effects of variations in lesson control and practice on learning from interactive video. Educational Communication and Technology Journal. 35(4), 203-212.

REFERENCES (Continued)

Johansen, K. J. & Tennyson R. D. (1983). Effect of adaptive advisement on perception in learner-controlled, computer-based instruction using a rule-learning task. Educational Communication and Technology Journal, 31(4), 226-236.

Kinzie, M. B., Sullivan, H. J. & Berdel, R. L. (1988). Learner control and achievement in science computer-assisted instruction. Journal of Educational Psychology, 80(3), 299-303.

Laurillard, D. M. (1984). Interactive video and the control of learning. Educational Technology. 26, 6, 7-15.

Levenson, P. & Morrow, J. (1986). A comparison of non-interactive and interactive video instruction about smokeless tobacco. Journal of Educational Technology Systems, 14, 3, 193-202.

Maw W. & Maw E. (1964). An Exploratory Study into the Measurement of Curiosity in Elementary School Children. Cooperative Research Project No. 801.

Maw W. & Maw E., (1965). Personal and Social Variables Differentiating Children With High and Low Curiosity. Newark, Delaware: University of Delaware.

Maw, W. & Maw E., (1966). Children's curiosity and parental attitudes. Journal of Marriage and the Family, 28, 343-345.

Merrill, M. D. (1975). Learner control. Beyond aptitude-treatment interactions. Audio-Visual Communications Review, 23, 217-226.

Parker, M. & Engel, J. (1983). A unified motivation and learning theory model. Education, 103(4), 353-360.

Phillips, T., Hannafin, M. & Tripp, S. (1988). The effects of practice and orienting activities on learning from interactive video. Educational Communication and Technology Journal, 36, 93-102.

Robson, E. H., Steward, A. P., & Whitfield, G. E. (1988). Pupils' choices in learning with computers. Journal of Computer Assisted Learning, 4, 93-102.

Ross, S. M. & Rakow, E. A. (1981). Learner control versus program control as adaptive strategies for selection of instructional support for math rules. Journal of Educational Pschology, 73, 745-753.

REFERENCES (Continued)

Santiago, R. & Okey, J. (1990, November). The effects of advisement and locus of control on achievement in learner-controlled instruction. Paper presented at ADCIS, San Diego, CA.

Schaffer, L. C. & Hannafin, M. J. (1986). The effects of progressively enriched interactivity on learning from interactive video. Educational Communication and Technology Journal, 34, 89-96.

Steinberg, E. R. (1989). Cognition and learner control: A literature review. 1977-1988. Journal of Computer-Based Instruction, 16(4), 117-121.

Tennyson, R. D. (1980). Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. Journal of Educational Psychology, 72, 525-532.

Tennyson, R. D. (1984). Artificial intelligence methods in computer-based instructional design: The Minnesota Adaptive Instructional System. Journal of Instructional Development, 3, 17-22.

Tennyson, R. D. & Buttrey, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. Educational Communication and Technology Journal, 28, 169-176.

Tennyson, R. D., Christensen, D. L., & Park, S. I. (1984). The Minnesota Adaptive Instructional System: An intelligent CBI system. Journal of Computer-Based Instruction, 11, 2-13.

Weintraub, S. (1986). Curiosity and motivation in scholarship. The Journal of General Education, 38, 3, 159-166.

Title:

Global Navigation in the Year 2001

Authors:

Ronald Aust
Jim Klayder

Global Navigation in the Year 2001

Ronald Aust and Jim Klayder
Instructional Technology Center
University of Kansas,

Paper presented at the 1990 Annual Conference of the
Association of Educational Communications and Technology

Near the turn of the 15th century, explorers who faced the unknown beyond their shores were anxiously awaiting the development of effective maps and navigational tools that would guide them safely across the oceans of our planet. As we prepare for the 21st century, the need for a new generation of navigational tools is evident. The quantity and global availability of information is increasing exponentially producing seas of data that are surging in both volume and diversity. For example, "A weekday edition of the *New York Times* contains more information than the average person was likely to come across in a lifetime in the seventeenth-century England (Wurman, 1989 p.6)."

These undulating waves of information content are being transformed further by rapid alterations in form. A myriad of display media, from video style newspapers to holographic arcade games, are emerging to convey information in new forms of animated diagrams, full motion video, "surround" sound and seemingly endless textual combinations of fonts, sizes, styles and orientations. Saul Wurman (1989) and others (Toffler 1970; Naisbitt, 1982) are concerned that instead of contributing to an era of enlightenment, this overabundance may result in mass afflictions of information anxiety and knowledge starvation. The greatest damage may come from being slow in responding to conspicuous need. Thus, the development of effective information management and consumption techniques will require some adjustments in the way we approach research designed to promote rapid development.

This paper is a report of the formative evaluation, insights and resulting modifications that have evolved during the initial phase of developing the Unified Network of Informatics in Teacher Education (UNITE). The objective is not to provide definitive data indicating that method A is more appropriate than method B. Rather, the intent is to participate in exchanging ideas for conducting research and developing information management tools that might lead to more effective strategies for benefiting from the vast array of information and media that we are facing as we prepare to enter the next century.

The Unified Network of Informatics for Teacher Education (UNITE)

The 1986 Holmes Group report explained that comprehensive approaches for integrating technology and education are needed to improve communication and collaboration among educators. UNITE was developed, with support from an Apple Corporation grant, to address the challenge of establishing a cooperative environment for exchanging ideas and computing resources among educators who have common goals but are separated by physical distances. Four constituent groups involved in the Teacher Education program are targeted: school administrators, teachers, interns (student teachers) and teacher educators. A primary objective is to investigate approaches to linking the network's administrative, methodological and instructional resources in ways that best reveal similarities among goals and ultimately foster a team spirit between the constituencies.

UNITE's hub is located in the Instructional Technology Center (ITC) at the University of Kansas School of Education. The hub consists of a primary distribution computer and a local area network tailored for the development of a range of educational resources including those that use emerging CD-ROM and interactive video technologies. The Hub's distribution computer uses standard telephone lines to connect to Macintosh stations in sixteen schools from six Kansas school districts serving over 83,000 students. The districts were selected on the basis of previous and projected clinical experience placements with consideration for equivalent grade level and minority representations.

Each station consists of: a Macintosh SE™ 20Meg hard drive, an Apple Modem™ and an ImageWriter II™ printer (Figure 1).

Prior to the beginning of the 1989 school year, the objectives of the project were described to district level administrators in each of the participating districts. These administrators selected schools that they believed would best serve the goals and objectives of the project. They also worked with the principals in the selected schools to identify the liaison team members.

Liaison teams for each school are comprised of four educators who assist ITC staff in integrating UNITE into their school's curriculum. This team consists of an administrator, a teacher facilitator, a librarian/resource coordinator, and a student intern from the School of Education. The liaison teams met with the project representatives to select the most appropriate site within their schools for placing the UNITE stations. Typically, the stations are located in the teachers work area or the school's library.

Initial design of the UNITE system software

During the spring of 1989, the instructional design team met to discuss alternatives for selecting or developing the UNITE software. Several commercially produced communication and navigation software packages were evaluated in terms of their ability to meet the criteria that were established according to the project objectives. The UNITE software was to consist of three primary components: electronic mail, a collection of educational resources and a Resource Navigator designed to serve as a "seamless" window to both the messaging system and the resources. The software must be easy-to-learn and easy-to-use because many of the target users are computer novices. The system must also provide meaningful support that the users would find useful in their daily educational activities.

After determining that existing software packages did not meet the criteria, the instructional design team established a plan for developing the software. The Resource Navigator would be developed first, including curriculum hierarchy for several subject areas. Model resources would then be selected and the links between the navigator and resources would be beta tested by students in the School of Education. Finally, the messaging system would be developed and beta tested. Upon later reflection, one of the best notions to come out of this initial instructional design phase was the decision to distribute the software earlier in beta versions instead of distributing it as a final 1.0 "fixed" version that would have a later release date and be more difficult to alter. The team believed that by doing this the resulting software would not only be more relevant to the needs of the users but it would also give the initial participants in the schools an awareness of participating in the development of the project resulting in a positive sense of ownership.

HyperText approaches to information navigation

A primary concern in developing the Resource Navigator was how to organize and provide access to the large and diverse information resources that were to be available through UNITE. Another issue was how to design an interface that encouraged the users to browse through resources that might be more regularly used by someone from outside their own constituency or content area. This second criteria was based on research of cooperative group learning (Johnson & Johnson, 1984,1986; Slavin, 1980) which has found that collaboration is more likely when group members have at least a passing awareness of the specialities and resources used by other members of the group. The instructional design team began by turning to existing examples of computer networks and literature on emerging HyperText strategies for accessing information.

Most existing systems used to locate resources available through computer networks (e.g. *CompuServe, The Source, Genie and BITNET*) are based on a command line interface. Searching for a resource generally begins with entering a user identification code followed by a password. The user is then "logged on" to the network, which means that they have made a connection through standard telephone lines, or possibly some other telecommunications medium, and are "talking to" a host computer at a remote sight. If a toll is imposed, the toll begins at the beginning of the search and lasts until the search is completed.

Once the user is logged on to the network, a code (sometimes a very cryptic code) is entered on a line that follows a flashing cursor to describe the parameter of the search. For example, to find resources related to the topic *Ecology* one might enter *%FS Ecology* and press the return key. The screen is then cleared and the first list of resources that relate to ecology appear on the screen. This can often be a very exhaustive list and the user must consult several screens to view the entire listing or select from a list of descriptors to further refine the search.

Although this type of searching can be efficient if the user knows exactly what s/he is searching for, the single line "search window" is very narrow and restrictive. Thus, the narrow search window of command line interfaces may limit the kind of "browsing through the stacks" that occurs when looking for a book in a conventional library. Often times this browsing results in locating books or other resources that expand horizons beyond the original intent of the library visit.

One alternative for expanding the likelihood of browsing is to adopt a HyperText approach for searches. Ted Nelson (1965, 1974, 1981, 1982) was the first person to use the term HyperText in reference to a multidimensional electronic notebook which allows users to access text in ways that support their thinking processes. The ideas for Nelson's term can be traced to Vannevar Bush (1945) who described a mechanism called the *memex* as, "A device in which an individual stores his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory." The visionary Douglas Englebardt introduced the first working example of the memex and the first mouse 20 years after Bush's description (Englebardt, 1963; Englebardt & Weeks, 1968). Both Bush and Englebardt believed that the human mind operates by association and that information should be presented accordingly. Their ideas have given rise to some more recent information navigation systems such as the *HyperText Abstract Machine (HAM)* (Campbell & Goodman, 1987), *Notecards* developed by a group at Xerox Park (Halasz, 1987) and the authoring environment *HyperCard* developed by Bill Atkinson at Apple Computer.

Development of the Educational Resource Navigator

After reviewing existing computer networks and HyperText literature the team decided to use HyperCard as the primary development environment for the UNITE software. The most recent version (1.0) of the Resource Navigator is represented in Figure 2. Subject areas are represented as buttons along the vertical axis of the screen. These subject categories can be cross-referenced to the support categories shown at the bottom of the screen. A science educator looking for a Computer Based Instruction (CBI) resource to support an ecology lesson would begin by clicking on the *Natural Science* button along the vertical axis and the *courseware* button along the horizontal axis. A window then appears showing the major sub-categories under Natural Sciences and the educator clicks on the biology button to see a second window overlap on the natural science window. The educator then clicks on the ecology button and sees a listing of computer-based courseware that is relevant to teaching about ecology. The user can then click on a resource (in this example a CBI simulation for the Apple II called *Oh Deer*) and see a screen or "card" that provides pertinent information about the *Oh Deer* simulation (Figure 3).

The information available on the resource catalog cards includes: a description of the courseware, objectives, intended grade level, a ratings by UNITE participants, information about the author of the courseware, equipment requirements and ideas for integrating the CBI. Users can also preview Macintosh CBI that is available on their station's hard drive.

Several revisions to the Resource Navigator were made prior to establishing the 1.0 version represented in Figure 2. A primary concern was how to organize and display a large amount of information within the limited screen size so that the users would be encouraged to browse through content areas and support categories. The original design had the navigator screen disappear once the content area was selected. For example, once *Natural Science* and *Courseware* were selected the screen would change so that only a listing of courseware pertinent to Natural Science was visible. The instructional design team later opted for the overlapping windows approach because they believed that it would encourage users to browse across both content and support categories.

The idea of overlapping windows can be traced to Allan Kay's work in the mid-seventies when he was with Xerox Park. The unique aspect of the UNITE Resource Navigator is that the overlapping windows scroll to the selected item and are fixed so that each time an item is selected it is visible at the top of the previous layer. This produces a layered hierarchy describing the path that the user has followed while at the same time providing visible and direct access to other information hierarchies.

The Resource Navigator is also designed with associative links that facilitate browsing across the support categories. For example, if the science educator viewing the "Oh Deer" card wishes to see a lesson plan that incorporates CBI on ecology, s/he can click on the *Teaching* button. Clicking on the *Research* button provides access to reports on the effectiveness of using CBI in *Science Education*, while clicking on the *Administration* button might provide certification criteria and state objectives for Science Education.

The Electronic Mail Component

All user input and output to the UNITE system occurs off-line. In other words, the information that is seen on the screen is being drawn from the stations hard drive and not from a host computer as is common in other network systems. When users sign on to stations they first see the navigation screen (Figure 2) where they can browse through the Resource Navigator or view mail by clicking on the mail box icon. When the mail box icon is clicked, a list of messages is seen, with the most recently received message appearing at the top of the screen. Once a user selects a message from the mail list, the screen depicted in Figure 3 appears. From this screen the user can continue to browse through messages, elect to write a new message, throw away "trash" messages, return to the Resource Navigator or sign off the system. Users can send messages to other individual users, to all users or to Special Interest Groups (SIGs).

UNITE is a multi-purpose distributed network. The main hub at the ITC is described as level 1, satellite hubs in long distance calling zones are level 2 and the user stations in the school are level 3. Once a message is written it resides only on that station's hard drive until the next connect is made. All computers in the system remain on at all times. The hub is in a continuous loop calling all level 3 stations within the local dialing area and periodically making a long distance call to the level 2 stations. Whenever a connect is made with a station, the hub sends all new messages it has received to that station. This transfer of messages may occur anytime during the day or night when the station is not being used.

Conducting formative evaluation through electronic mail

Formative evaluation has been defined as "a process of systematically trying out instructional materials with learners in order to gather information and data which will be used to revise the materials (Dick, 1977 p. 311)." There is wide acceptance in the education community that formative evaluation is an important step in the instructional design process. Little research has been done, however, on the effectiveness of formative evaluation. The few studies that have been done (Robeck, 1965; Baker, 1970; Komoski, 1974) principally indicate that instructional design which is revised on the basis of formative evaluation has significantly better learning outcomes than the original designs.

The formative evaluation procedures roughly approximated Dick's three stage method for evaluating instructional software. These stages include: one-to-one evaluations (one to three students); small-group evaluations (five to 15 students); and field tests using 20 or more subjects (Golas, 1983).

A continuous loop of evaluation and revision was used during the four evolutions of the "pre-release" system software. The first version of the system (0.6) was evaluated with a one-to-one technique. Suitable revisions were made based on the first round of evaluation and the software was placed on the school stations as version 0.7.

The message system played a critical role in providing feedback during stage two and stage three evaluations of the early (0.7, 0.8, and 0.9) versions of the software. Prior to the release of version 0.7 two Critic's SIG's were established. One Critic's SIG was used when making suggestions regarding the message system, the other was for suggesting changes to the Navigator and Resource Catalogs. During orientation sessions in the schools, the UNITE staff explained to the liaison teams that the system must eventually be easy-to-learn, easy-to-use and intuitive so that teachers could rapidly obtain available support before class or during their planning period. The importance of the Critic's SIG was explained and participants were encouraged to make forthright and frequent recommendations.

Over 350 messages were sent to the Critics SIGs. They included suggestions: for adding new subject hierarchies, changes in design and placement of icons, procedures for accessing and organizing the mail system, changes to the step-by-step procedure for enrolling new users and revisions to the on-line instructions. The messages from the Critic's SIGs were very useful in making upgrades to the system software. As opposed to the interview sessions, that were also conducted during this phase of evaluation, the Critic's messages seemed to be more direct reflections of user concerns. Possibly this is because the users were allowed to evaluate the systems according to their own schedule and within the environment in which the final system would be used. The Critic's SIGs are readily available whenever the user encounters a problem and wishes to make a comment. For example, when a user comes across a portion of the subject hierarchy within the navigator that is confusing s/he simply clicks the mail box icon, sends a message to the Critic's SIG and then returns to Resource Navigator to continue browsing.

The use of electronic messages during formative evaluation is a practical means of gauging effectiveness across representative samples of a user population that is separated by physical distances. Although the UNITE system is being evaluated in the relatively confined area of eastern Kansas, it could be extended to English speaking populations across the globe and the Critic's SIGs would continue to function to make the system relevant to the needs of an even broader population. The use of Critic's SIG messages also allows the software developers to be more responsive than would be possible using direct observations or face-to-face interviews. The collection of evaluations is more immediate and some programming changes can be made by sending special "system" messages that alter the software residing on the hard drives in the remote sites.

Community-published dynamic resources

One advantage of distributed systems is that all resources are viewed at machine speed and displays are limited only by the local system software. Thus, more sophisticated and easier to use resources with graphical interfaces can be developed. Some of the more well received UNITE resources have been produced in the ITC to support regional educational activities such as a catalog of field trips activities. When several of the Critic's SIG messages referred to an interest in contributing to these resources, the developer responsible for the Field Trip catalog sent a message to all users asking for suggested additions or modifications to the catalog (the message seen in Figure 4 is a response to one of these requests). In our initial solution for adding user contributions to the resources, the suggested changes to the Field Trip catalog were made at one of the computers in the ITC's local area network. A new version of the Field Trip catalog was then manually transferred to the station's hard drives with each update to the system software.

We were especially interested in refining this contribution strategy after reviewing the literature on user participation in technological innovation. Larry Cuban (1986) has provided evidence that teachers will most often resist educational innovation when they believe that the innovation is being pushed on them. Robert Heinich (1984) has found that teachers will go so far as to intentionally sabotage new forms of technology if they feel threatened. However, a recent study that we conducted on *Teacher Attitude Toward Instructional Technology*, (Aust et al., 1989) revealed that teachers were excited about the prospect of being involved in technological innovation provided that they could be involved in shaping the innovation. This was one of reason for releasing the system software in a beta version and the impetus for the development of community-published dynamic resources.

A more recent solution to the problem of how to add user contributions to the resources is represented by the icon of the hand holding the light bulb seen in Figure 2 of the CBI catalog. When a user wishes to add a record to the catalog of CBI resources s/he clicks on this *contribution* button icon and is given the option of either adding a new record or modifying an existing record. Once the *Add a New Record* option is selected, an entry screen appears and the user is prompted through a step-by-step series of questions for entering the new record. This process is similar, although somewhat more involved, to the procedure used when withdrawing money from a cash machine at a bank. For example, the user is prompted to enter: the title of the CBI, information about the publisher and vendor, the equipment requirements, a description of intended users, and a brief description of the CBI. The process for modifying a record is similar to adding a record, with the exception that all information in the record that is being modified appears on the entry screen.

After the user is satisfied that all information for the contribution has been entered properly, a *send* button is clicked and an encoded message representing the entry is placed in a holding pen waiting for the next connect with the the hub. Message for resource entry received at the hub are compiled in an *Evaluator's Stack* and reviewed for content and accuracy. Once approved, a *change resource* system message is sent to the remote stations where the approved entry is appended to the appropriate resource. A *change catalog* system message can also be issued from the hub to delete records.

By using this approach, contributions to resources on the hard drives can be made by any user on the UNITE system, in other words, they are community publications by the users. These resources are also dynamic in the sense that they are going through metamorphoses as influenced by both user contributions and usage patterns. Because these community-published dynamic resources reside on the hard drives of the user's computer, they also operate much faster and include graphical interfaces that are not available when viewing interfaces to long distance networks that are reading data directly from a host computer. The community-published dynamic resources that have been developed thus far include: the CBI catalog, a Lesson Plan catalog, Field Trips, and an index of abstracts from educational research/method articles.

We see many promising possibilities for these community-published dynamic resources. There is the prospect of dynamically constructing catalogs of exemplary Lesson Plans produced by teachers from throughout the country or world. We now have the ability to quickly disseminate resources dealing with timely topics such as drug or AIDS education in a consistent format. Perhaps the most important implication may have to do with the positive social and psychological impact that community-published dynamic resources could have on educators. By contributing to these resources, the user becomes an active participant in technological innovation not simply a passive bystander caught in waves of information and technological overload. Collaboration and cooperation among educators who have common goals might also be enhanced as they mutually contribute to the development of community-published dynamic resources.

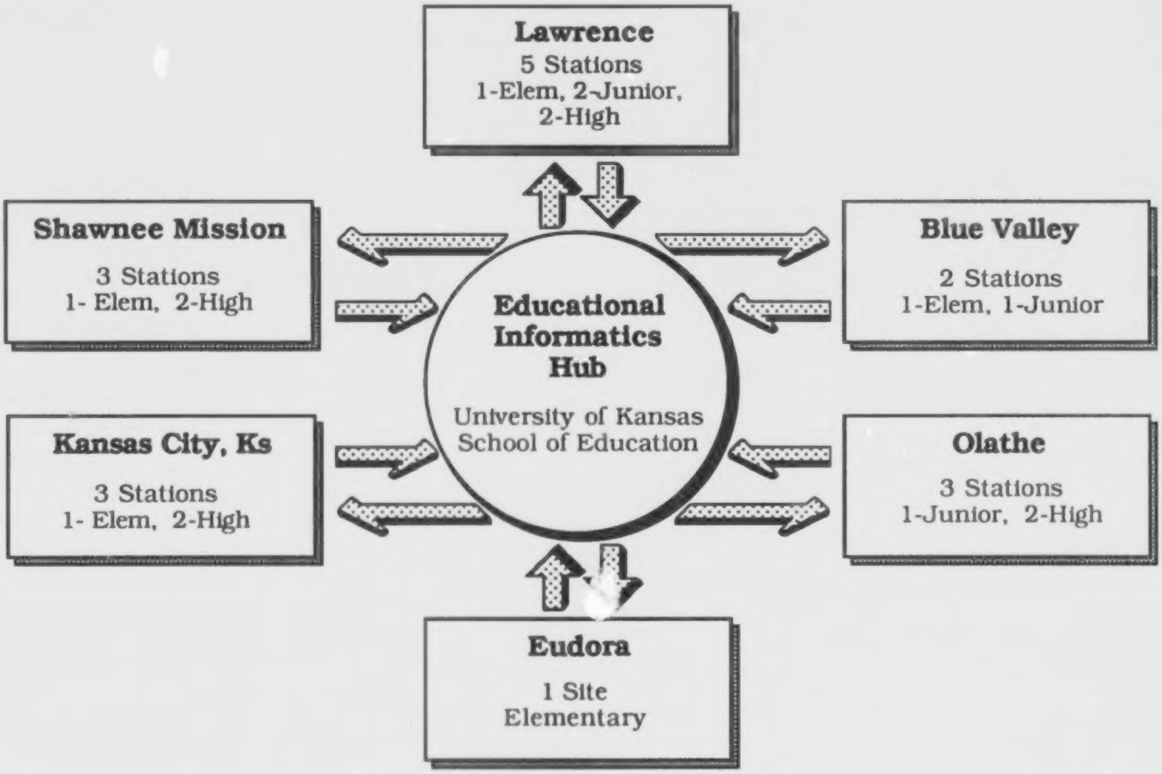
Developing global participation

As we approach the 21st century, more effective means for navigating through the ever expanding seas of information are needed. Computers appear to be the tools of the information age and as Ted Nelson prophesied, "If computers are the wave of the future then displays are the surfboards." Icons and graphical interfaces are clear contributions to presenting information about processes and relationships in a concise manner. But much remains to be done before their potential will be realized. Standards will need to be established in such areas as: the use of windows, pull-down menus, animation for transitions between information resources, hierarchies for organizing information and universally understood icons that cross language barriers.

To meet this challenge, formative evaluation should be conducted across great distances and in a short amount of time so that strategies can be developed prior to the next wave. Electronic networks may serve as useful tools both for gathering evaluations and conducting rapid modifications to systems designed for organizing and managing information. These development strategies must also consider the social and psychological impact of adjusting to the Information Age. One possible remedy is to provide community-published resources that are relevant to user needs and allow educators to participate in the innovations of the 21st Century.

Figure 1

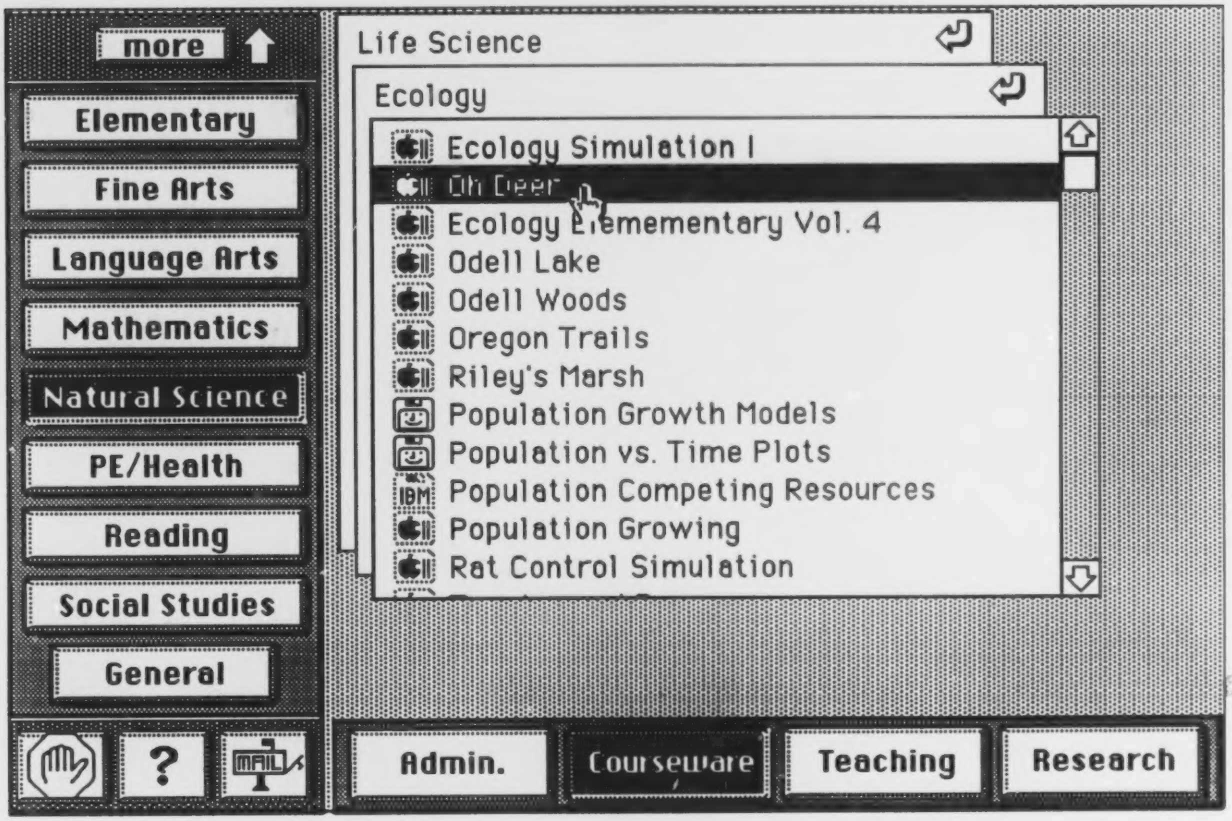
Schools who participated in the UNITE project


Lawrence
5 Stations
1-Elem, 2-Junior,
2-High
Shawnee Mission
3 Stations
1- Elem, 2-High
Blue Valley
2 Stations
1-Elem, 1-Junior
Educational Informatics Hub
University of Kansas
School of Education
Kansas City, Ks
3 Stations
1- Elem, 2-High
Olathe
3 Stations
1-Junior, 2-High
Eudora
1 Site
Elementary

The 17 stations were selected on the basis of previous and projected clinical experience placements with consideration for equivalent grade level and minority representations. Each station includes: 1 Macintosh SE™ 20Meg hard drive, 1 Apple Modem and 1 ImageWriter II™.

Figure 2

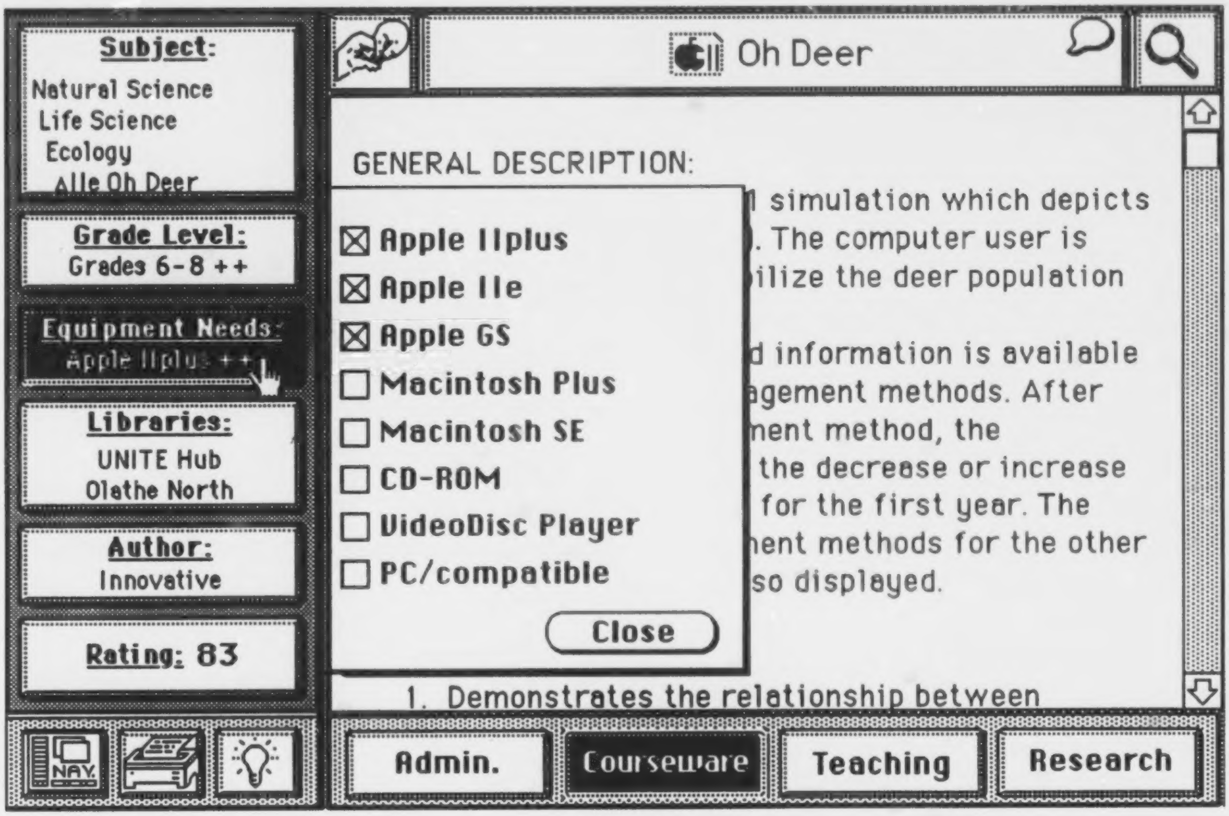
UNITE Resource Navigation System


more
Elementary
Fine Arts
Language Arts
Mathematics
Natural Science
PE/Health
Reading
Social Studies
General
?
MAIL
Life Science
Ecology
Ecology Simulation I
Oh Deer
Ecology Elemementary Vol. 4
Odell Lake
Odell Woods
Oregon Trails
Riley's Marsh
Population Growth Models
Population vs. Time Plots
IBM
Population Competing Resources
Population Growing
Rat Control Simulation
Admin.
Courseware
Teaching
Research

Content categories are located vertically along the left margin of the screen and the four support categories, *Administration*, *Courseware*, *Teaching* and *Research* are located along horizontally along the the bottom of the UNITE resource navigator. In this view the user began by selecting the support category for *Courseware*. The main content category *Natural Science* and then the sub-categories of *Life Science* and *Ecology* were selected. The user has just selected the software the software Oh Deer which is an ecological simulation of an overpopulated deer heard. The ΔIIe that precedes the courseware name indicates that the *Oh Deer* CBI is for an Apple IIe computer, that it is available in at least one of the participating schools and that there is an entry for *Oh Deer* in the UNITE catalog of courseware. Having clicked on the *Oh Deer* courseware the user will be taken directly to the catalog entry depicted in Figure 3.

Figure 3

A UNITE screen showing an entry for the *Oh Deer* Computer-Based Instruction Catalog



Subject:
Natural Science
Life Science
Ecology
Alle Oh Deer

Grade Level: Grades 6-8 ++

Equipment Needs: Apple IIplus ++

Libraries:
UNITE Hub
Olathe North

Author: Innovative

Rating: 83

Oh Deer

GENERAL DESCRIPTION:

l simulation which depicts
). The computer user is
ilize the deer population
d information is available
agement methods. After
ment method, the
the decrease or increase
for the first year. The
ment methods for the other
so displayed.

1. Demonstrates the relationship between

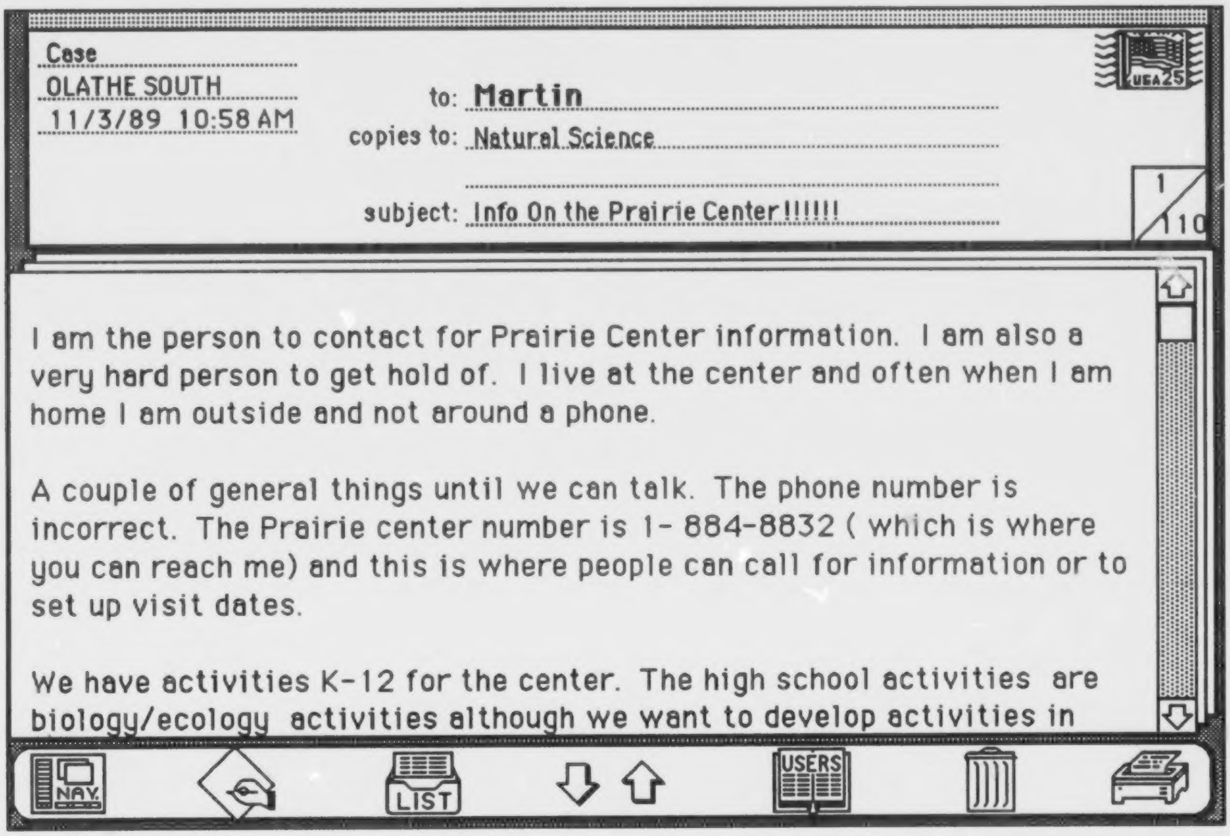
☒ Apple IIplus
☒ Apple IIe
☒ Apple GS
☐ Macintosh Plus
☐ Macintosh SE
☐ CD-ROM
☐ VideoDisc Player
☐ PC/compatible

This screen makes available all basic information related to a piece of courseware. The scrolling field in the background contains a description of the courseware and suggestions for its use. The buttons in the left margin are pop up displays of detailed information. The user is currently viewing the *equipment needs* for *Oh Deer*.

The format of the catalog of lesson plans is also similar to this screen with the following information available through the buttons in the left margin; subject, activities, intended learners, objectives, resources and author. The research catalog includes: title, reference, subject, descriptors, abstract, reviewers and libraries.

Figure 4

A UNITE Messaging System



Case
OLATHE SOUTH
11/3/89 10:58 AM

to: Martin
copies to: Natural Science
subject: Info On the Prairie Center!!!!!!

1/110

I am the person to contact for Prairie Center information. I am also a very hard person to get hold of. I live at the center and often when I am home I am outside and not around a phone.

A couple of general things until we can talk. The phone number is incorrect. The Prairie center number is 1- 884-8832 (which is where you can reach me) and this is where people can call for information or to set up visit dates.

We have activities K-12 for the center. The high school activities are biology/ecology activities although we want to develop activities in

This actual message was sent over the network on November 3, 1989. The upper right corner shows a cancelled stamp indicating that the message has been mailed. Directly below the stamp are numbers showing that this is message 1 of 110. From left to right along the bottom of the screen are:

1. the button that returns the user to the Resource Navigator;
2. the button for writing a message;
3. the mail list button;
4. the arrow button for going to the next message
5. the arrow button for going to the previous message;
6. the button that shows additional information about the users, including: the content and grade level they teach, where they teach and their special interest groups;
7. the trash can button for throwing away messages;
8. the printer button .

References

Aust, R., Allen G., Bichelmeyer, B. (1989). Integrating instructional technology in educational institutions: the proper role for teachers. *Proceedings of selected research papers 1989 convention AECT*, Instructional Resources Center, Iowa State : Ames.

Baker, E.L. Generalizability of Rules for an Empirical Revision. AV Communication Review, 1970, 18(3), 300-305.

Bloom, A. (1987). The closing of the american mind. New York: Simon and Shuster.

Branch C.E., Ledford, B.R., Robertson, B.T. & Robinson, L. (1987) The validation of an interactive videodisc as an alternative to traditional teaching technique: auscultation of the heart Educational Technology, 16-22

Bruder, I. (1989). Future teachers: Are they prepared? Electronic Learning, *8,(4)* 32-39.

Bush, Vannevar. (1945). As we may think. Atlantic Monthly, 176, 1 (July 1945): 101-108.

Bush, Vannevar. (1967). Science is not enough, Vannevar Bush. William Morrow.

Campbell, B. & Goodman, J. (1988) HAM: General-Purpose hypertext abstract machine. Communications of the Association of Computing Machinery, 31(7).

Caron, J., Jamieson, D. & Dewar (1980). Evaluating pictograms using semantic differential and classification techniques. Ergonomics, 23, 137-146.

Cuban, L. (1986). Teachers and machines. New York: Teachers College Press.

Dick, W. (1977). Formative Evaluation. In L. Briggs (Ed.), Instructional Design: Principles and Applications (pp. 311-333). Englewood Cliffs, NJ: Educational Technology Publications, Inc.

Engelbart, D. C. & English W. K. (1968). A research center for augmenting human intellect. AFIPS Proceedings, Fall Joint Computer Conference.

Engelbart, D. C. (1963) A conceptual framework for the augmentation of man's intellect. in P. D. Howerton, and D. C. Weeks (Eds), Information Handling (vol. 1), Washington, D.C.: Spartan Books

Frymier, J. (1987). Bureaucracy and the neutering of teachers. Kappan, *69*, 9-15.

Gittins, D. (1986). Icon-based human-computer interaction. International Journal of Man-Machine Studies, 24, 519-543.

Golas, K. (1983). The Formative Evaluation of Computer-Assisted Instruction. Educational Technology, 23(1), 26-28.

Grady, D. (1988). Giving teachers their due. Kappan *70*, 31.

Grossman, D. (1987). Electronic college courses: The professor must be in charge. The Chronicle of Higher Education 2/11.

Guba, E. & Lincoln, Y. (1981). Effective Evaluation. San Francisco, CA: Jossey-Bass.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. Educational Communications and Technology Journal, *29*, 75-91.

Halasz, F. Reflections on notecards: Seven issues for the next generation of hypermedia systems. Communications of the Association of Computing Machinery, 31(7)

Heinich, R. (1984). The proper study of instructional technology. Educational Communications and Technology Journal, *32*, 67-87.

Johnson, D. W., Johnson, R. T., Holubec, E. J., & Roy, P. (1984). Circles of learning: Cooperation in the classroom. Minneapolis, MN: Association for Supervision and Curriculum Development.

Kennedy, J. (1974). Icons and Information. In D. Olson (Ed.), Media and Symbols: The Forms of Expression, Communication, and Education (pp. 211-240). Chicago, IL: University of Chicago Press.

Komoski, P.K. An Imbalance of Product Quantity and Instructional Quality: The Imperative of Empiricism. AV Communication Review, 1974, 22(4), 357-386.

Leininger, M.M. (1985). Qualitative research methods in nursing. New York, NY: Grune & Stratton.

Lepper, M. & Gurtner, J.L. (1989) Children and computers. American Psychologists 44(2) 170-178.

Lincoln, Y. & Guba, E. (1985). Naturalistic Inquiry. Beverly Hills, CA: Sage.

Lutus, P. (1985) Computer math. Popular computing March 103-106.

Maeroff, G.I. (1988). A blueprint for empowering teachers. Kappan *69*, 472-476.

Marcus, A., (1984). Graphic design for computer graphics. IEEE Computer Graphics and Applications, 3, 63-68.

Mischler, E., (1978). Meaning in context: Is there any other kind? Educational Review, 19(1), 1-19.

Naisbitt, J. (1982). Megatrends. New York, NY: Warner Books, Inc.

Nelson, T. H. (1965). The hypertext. Proceedings International Documentation Federation Annual Conference.

Nelson, T. H. (1974). Dream machines: New freedoms through computer screens - A minority report. Computer Lib: You Can and Must Understand Computers Now. Chicago, IL: Hugo's Book Service.

Nelson, T. H. (1981) Literary machines. Swarthmore, PA: T.H. Nelson.

Nelson, T. H. (1982) A new home for the mind. Datamation (March, 1982): 169-180.

Patterson, A. & Bloch, B. (1987). Formative Evaluation: A Process Required in Computer-Assisted Instruction. Educational Technology. 27(11), 26-30.

Reed, S. (1985). Effect of computer graphics on improving estimates to algebra word problems. Journal of Educational Psychology 77(3). 285-298.

Rhodes, L. (1988). We have met the system-and it is us. Phi Delta Kappan 70 (1), 28-30.

Ridley W.J. & Hull M.H. (1986) Transforming American Education: Reducing the Nation at Risk to the Nation. A report to the Secretary of Education, United States Department of Education, by the national Task Force on Educational Technology, Washington D.C.: Office of Educational Research and Improvement.

Robeck, M.D. A Study of the Revision Process in Programmed Instruction. Unpublished Master's Thesis, University of California, Los Angeles, 1965.

Rossenberg, S. & Bitter, G. (1988). Microcomputer infusion project: A model. Tech Trends *33*, (5),

Slavin, R.E. (1980). Cooperative learning. Review of Educational Research, 50, 315-352.

Smith, J. B. & Weiss, S. F. (1988) HyperText: An overview. Communications of the Association of Computing Machinery, 31(7).

teacher development. Electronic Learning 8(7) 22-24.

Tofler, A. (1970). Future Shock. New York, NY: Random House.

Watt, D (1989) Research on Logo learning as a path to professional

Wurman, R. (1989). Information Anxiety. New York, NY: Doubleday.

Title:

The Interaction of Color Realism and Pictorial Recall Memory

Author:

Louis H. Berry

THE INTERACTION OF COLOR REALISM
AND
PICTORIAL RECALL MEMORY

Louis H. Berry
University of Pittsburgh
Pittsburgh, PA

During the past several years, extensive research has addressed the effects of variations in color realism on visual recognition memory. Few studies have focused on the effects of these color factors on pictorial recall memory. It was the purpose of this study, therefore, to investigate this interaction to: 1. better understand the effects of variations in color realism 2. to draw comparisons between visual recall memory and visual recognition memory in terms of color information processing.

A sizable mass of research has focused on the variable of color in instruction (Dwyer 1972, 1978, 1987; Berry, 1974; Winn, 1976; Chute, 1979; Lamberski, 1980). These investigations addressed one aspect of the larger theoretical debate regarding visual complexity and human information processing. It has long been contended that the mere addition of visual cues will increase the ability of the viewer to store and retrieve visual information.. This orientation, termed "realism theory" by Dwyer (1967), has strong theoretical foundations (Dale, 1946; Morris, 1946; Carpenter, 1953 and Gibson, 1954) and was the notion on which cue summation theory is predicated (Severin, 1967). Other theorists and researchers (Broadbent, 1958, 1965; and Travers, 1964) have argued against this theoretic base on the grounds that the human information processing system is limited in capacity and in times of rapid information reception, irrelevant information may block the processing of other, relevant information. Studies (Kanner, 1968; Katzman and Nyenhuis, 1972; Dwyer, 1972, 1978, 1987) have investigated this apparent inconsistency with contradictory results.

The inclusion or absence of color information can be regarded as one dimension of visual complexity. Color can function in a dual role when used in visual displays. First, it can serve primarily a coding function, providing additional information but not providing any realistic description of the display. In this case, the effectiveness of color can be predicted by cue summation theory, but not by the realism hypothesis. Alternately, color can be used to present a more realistic version of the visual display. In this instance, in addition to providing a greater number of overall cues, it provides the viewer with more realistic attributes or "handles" with which to store and retrieve information. When color is used in this fashion, its value could be predicted by the realism theories as well as by cue summation theory.

Much past research investigating the differences between color and monochrome visuals failed to take into account the fact that realistic color visuals contain intrinsically more information and consequently require more time for processing. In an attempt to resolve this methodological inconsistency as well as to more accurately assess the role of color in human information processing, Berry (1974) compared realistic and non-realistic color versions of the instructional materials on the human heart developed by Dwyer (1967). Data suggested that, in those learning tasks where visual materials contributed significantly to the improvement of instruction, realistic color materials were most effective. Later research (Berry, 1977, 1982, 1983, 1990) which investigated the color realism/complexity question relative to pictorial recognition memory found both realistic and non-realistic color materials superior to monochrome visuals. These findings suggest that cue summation theory may provide an accurate description of how color functions in basic information processing tasks such as picture recognition.

More limited research has addressed the effects of visual complexity on recall memory. Ritchey (1982) reported an advantage in recall for outline drawings over detailed drawings. A study conducted by Jesky (1984) suggested the superiority of color over black & white and both color and black & white visuals respectively over line drawing images in a recall task. These findings were further confirmed by Alfahad, 1990. Recall memory involving the specific variable of color realism has, however, not been investigated.

METHOD

The stimulus materials used in this study were modified versions of those developed by Jesky (1984). The original materials consisted of three sets of visuals, each produced in three visual formats: line drawing, black & white and color. To create the sets of visuals, three different collections of common household items (32 per set) were randomly arranged on a neutral photo backdrop. In selecting the objects, care was taken to ensure that no verbal labels, names or symbols were visible. Each set was photographed on color slides and then later recopied onto black and white slides. A line drawing of each scene was traced by an artist working from the projected black and white slides. The resulting drawings were also copied onto 35mm slides. In the present study these materials were modified by the creation of a nonrealistic version for each set. This was achieved by photographically reversing the color (realistic) slides of each set so that each color was systematically reversed to it's complementary value.

The color reversal process permitted the overall number of color cues to be held constant, while the degree of color realism was manipulated. In this case, the color cues no longer represented meaningful or familiar attributes of the images, but rather irrelevant visual information. For the purposes of this study, the complete group of materials consisted of three different sets of slides, each containing identical images produced in four different visual formats: line-drawing, monochrome (black & white, realistic color and nonrealistic color.

For presentation purposes, the stimulus materials were organized in a 3 x 4 grid and rotated so that each visual format in each set could be presented to different groups of subjects as shown in table 1.

Table 1
Presentation order of the visual treatments
and stimulus sets to the subject groups

	Line Drawing	Black/ White	Realistic Color	Nonrealistic Color
Group 1	Set 1	Set 2	Set 3	—
Group 2	—	Set 1	Set 2	Set 3
Group 3	Set 3	—	Set 1	Set 2
Group 4	Set 2	Set 3	—	Set 1

The research sample consisted of forty students enrolled in graduate and postbaccalaureate courses in education. All subjects were consenting volunteers.and were randomly assigned to one of the four treatment groups.

A free recall procedure was employed in which subjects were presented with the stimulus materials for a brief study time after which they were asked to recall as many items as they could from the stimulus images. The limited capacity of short term memory and rehearsal difficulty predicts that within a brief period of time (approximately four minutes), all items available in memory will have been recalled.

The stimulus materials were presented by means of a Kodak Ektagraphic Carousel slide projector in a semi-darkened room. Projected image size was maintained at four by six feet and subjects were seated in two rows between eight and fourteen feet from the screen. Each subject group viewed an individual slide for a period of 20 seconds after which the slide was removed from view and the lights raised. Subjects then received four minutes to write down as many objects as they could recall from the slide. Prior research, (Salomon & Cohen, 1977) indicated that an exposure time of 20 seconds and a response period of four minutes was adequate time for all subjects to complete a recall task similar to that employed in this study. This procedure was repeated again with each of the other two slides assigned to the particular group of subjects with a rest period of five minutes provided between each slide A similar procedure was followed with each of the other three groups of subjects.

ANALYSIS

Data obtained from each experimental group was rotated and collapsed under the four respective treatments for purposes of analysis. Although the data could not be analyzed by means of a repeated measures design where every subject receives every treatment, a high degree of redundancy was achieved by means of the rotation procedure. Means and standard deviations for each of the treatment groups are shown in Table 2.

Table 2.
Means and Standard Deviations for Color Treatment groups

Treatment	Mean	S.D.
Line drawing	7.90	1.87
Black & White	9.75	1.72
Realistic color	10.68	2.02
Nonrealistic Color	6.50	1.74

A one-way analysis of variance was performed on the data, resulting in a significant F value (F=41.31, p<.0001). Post hoc comparisons via the Scheffé method revealed a number of significant pair-wise comparisons. These differences are presented in Table 3.

Table 3.
Summary of Pair-wise Comparisons for Color Treatments

Black & White > Line Drawing
Realistic color > Line Drawing
Line Drawing > Nonrealistic Color
Black & White > Nonrealistic Color
Realistic Color > Nonrealistic Color

DISCUSSION

Data obtained confirm the findings of Jesky (1984) for the line drawing, black & white and realistic color treatments. The additional treatment, nonrealistic color was significantly lower than all other treatments. This result would indicate that the nonrealistically colored materials were substantially more difficult to recall. Examination of the data demonstrates that as the variable of visual complexity increases, so does the degree of recall. In this case, the most realistic materials (realistic color) were most effective in facilitating visual recall and the least complex (line drawing) were least effective. The nonrealistic color treatment was systematically as visually complex as the realistic color treatment in the total number of visual cues, but did differ in terms of realistic color referents. In this case it can be concluded that the use of nonrealistic color did not aid recall by supplying additional cues, but rather interfered with the recall task by providing irrelevant, possibly distracting information. From this finding, it seems apparent that the factor of visual realism contributes substantially to the recall process by providing familiar, functional cues which aid the learner in retrieving the information. This would seem to imply that the learner has a more generalized schema of the objects which incorporate familiar color codes. When realistic colors are used in the encoding process, these color codes help in retrieving the schemata. In the instance of nonrealistic color, the unfamiliar color cues may inhibit retrieval of the image because they conflict with those which the viewer expects or has instantiated into their cognitive structure.

These findings tend support other research which investigated the effectiveness of realistic and nonrealistic color on a pictorial recognition task (Berry 1982, 1990). In a recognition task, the nonrealistic color cues perform a useful task in providing relevant, although unfamiliar codes used by the viewer in both storage and retrieval processes. In the recall process, the images may be stored based upon shape or verbal labels rather than overall visual cues and consequently are retrieved via a more general schema which does not rely on the unique color cues. To the contrary, these cues may interfere with the process of incorporating the new images into an existing knowledge structure.

The results of this study also lend credence to the "realism theory" orientation, but do not support the generalized theory of cue summation which predicts that learning will be increased as the total number of cues increases. In the case of nonrealistic color, the simple addition of the nonreal cues inhibits rather than facilitates recall of the information.

REFERENCES

Alfahad, F. N., (1990). The interactive effects of cognitive functioning and visual realism on visual memory recall tasks. (Unpublished doctoral dissertation, University of Pittsburgh).

Berry, L.H. (1974). An exploratory study of the relative effectiveness of realistic and non-realistic color in visual instructional materials. (Doctoral dissertation, the Pennsylvania State University). *Dissertation Abstracts International, 35,* 7717a.

Berry, L.H. (1977). *The effects of color realism on pictorial recognition memory*. Paper presented at the annual convention of the Association for Educational Communications and Technology, Miami Beach, FL.

Berry, L.H. (1982). Signal detection analysis of color realism data. *Instructional Communications and Technology Research Reports. 13,*

Berry, L. H. (1983). *The effect of cross cultural variations and color realism on pictorial recognition memory*. Paper presented at the annual convention of the Association for Educational Communications and Technology, New Orleans, LA.

Berry, L. H., (1990). Effects of hemispheric laterality on color information processing. *Perceptual and Motor Skills. 71,* 987-993.

Broadbent, D.E. (1958). *Perception and Communication*. New York: Pergamon Press.

Broadbent, D.E. (1965). Information processing in the nervous system.*Science. 35,* 457-462.

Carpenter, C.R. (1953). A theoretical orientation for instructional film research. *AV Communication Review, 1,* 38-52.

Chute, A.G. (1979). Analysis of the instructional functions of color and monochrome cueing in media presentations. *Educational Communications and Technology Journal. 27,* 251-263.

Dale, E. (1946). *Audio-visual methods in Teaching*. New York: Dryden Press.

Dwyer, F.M. (1967). Adapting visual illustrations for effective learning. *Harvard Educational Review,.37,* 250-263.

Dwyer, F.M. (1972). *A guide for improving visualized instruction..* State College, PA: Learning Services.

Dwyer, F.M. (1978). *Strategies for improving visual learning*. State College, PA: Learning Services.

Dwyer, F. M., (Ed.).(1987) *Enhancing Visualized Instruction - Recommendations for Practice* . State College: Learning Services.

Gibson, J.J. (1954). A theory of pictorial perception. *A V Communications Review. 2,* 2-23.

Jesky, R. (1984). The interactive effects of pictorial presentation and cognitive style on a visual recall memory task (Doctoral dissertation, University of Pittsburgh, 1985). *Dissertation Abstracts International, 44,* 11077.

Kanner, J.H. (1969). *The instructional effectiveness of color in television: A review of the evidence.* Palo Alto, CA: Stanford University.

Katzman, H. & Nyenhuis, J. (1970). Color vs black and white effects on learning, opinion and attention. *AV Communication Review. 20,* 16-28.

Lamberski, R.J. (1980). *A comprehensive and critical review of the methodology and findings in color investigations.* Paper presented at the annual convention of the Association for Educational Communications and Technology, Denver, CO.

Morris, C.W. (1946). *Signs, language and behavior.* New York: Prentice-Hall.

Ritchey, G. H. (1982). Notes, comments and findings: Pictorial detail and recall in adults and children. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 2, 139-141.

Severin, W. (1967). Another look at cue summation. *AV Communication Review. 15,* 233-245.

Salomon,, G. & Cohen, A. A., (1977). Television formats, mastery of mental skills, and the acquisition of knowledge. *Journal of Educational Psychology. 69.* 612-619.

Travers, R.M.G. (1964). The transmission of information to human receivers. *AV Communication Review. 12,* 373-385.

Winn, W.D. (1976). The structure of multiple free associations towards black and white pictures and color pictures. *AV Communications Review. 24,* 273-293.

Title:

Visual Complexity and Pictorial Memory: A Fifteen Year Research Perspective

Author:

Louis H. Berry

VISUAL COMPLEXITY AND PICTORIAL MEMORY: A FIFTEEN YEAR RESEARCH PERSPECTIVE

Louis H. Berry
University of Pittsburgh
Pittsburgh, PA

During the past fifteen years, an ongoing research project at the University of Pittsburgh has focused on the effects of variations in visual complexity and color on the storage and retrieval of visual information by learners. The purpose of this paper is to describe and summarize this group of studies and to draw conclusions from the research which may more clearly define the role of color and its application to the design of visual instructional materials.

Research has clearly indicated that visual materials facilitate instruction, but research has not fully delineated how the design factors of visual complexity, and color, specifically, interact with cognitive processes to produce this advantage. Over the years, a number of theoretic positions have addressed the question of visual complexity and realism from somewhat conflicting directions.

One position, termed "realism theories" by Dwyer, (1967) suggests that instructional materials become more effective as the degree of realism or the overall number of visual and contextual cues is increased. This hypothesis, often referred to as cue summation theory (Severin, 1967), was a popular component of many theories put forth during the 1940's and 1950's, (Dale, 1946; Morris,1946; Carpenter,1953; Gibson, 1954). A closer evaluation of these conceptualizations, however indicates two distinct interpretations. Initially, cue summation theory describes learning being enhanced as the total number of visual cues increases. No distinction is made between those which contribute to a more realistic portrayal of the information and those cues which are simply added to the display for enhancement purposes, such as color codes or stylized color for aesthetic appeal. In the case of color as a realistic descriptor, the cues contribute additional elaborative or meaningful information which may or may not be useful in encoding the visual information in memory. The simple addition of irrelevant visual information does not necessarily insure that the cues will be used by the viewer for memory storage unless a complex coding scheme is previously constructed. In either case, the method whereby the cue information is used in memory storage is still unclear. A conflicting set of theoretic orientations (Broadbent, 1958, 1965; Travers, 1964) suggested that the human information processing system is of limited capacity, and in times of high information processing, irrelevant cues may effectively block the processing of other, relevant information. These "filter" or information overload theories have been modified through research over the past years; however they remain highly credible descriptors of the processing system because they are supported by the large base of research relative to the limited nature of human information processing. Both of these theoretical positions received substantial research attention and became the focus of a large number of, albeit conflicting, research investigations. Studies conducted by Dwyer and others (1967, 1972, 1978, 1987) addressed the specific factors of complexity and color in visualized learning in a comprehensive and systematic manner The findings suggested that the complexity/color question is strongly dependent upon the specific instructional objective and learning task. Other research reported by Dwyer (1987) addressed additional factors among which were cognitive style, retention, visual realism and learner aptitudes. Additional research conducted over a period of many years has addressed the issue of the relative effectiveness of color in instructional materials. Studies, (Vandermeer, 1952; Kanner, 1968; Katzman & Nyenhuis,1972; Winn, 1976; Chute, 1979; Lamberski, 1980), have investigated this controversial topic with conflicting results.

The inclusion or absence of color information can be regarded as one dimension of visual complexity. Color can function in a dual role when used in visual displays. First it can serve primarily a coding function, providing additional information, but not providing any realistic description of the elements of the display. In this case, the effectiveness of color can be predicted by cue summation theory, but not by the realism hypothesis. Second, color can be used to present a more realistic version of the visual display. In this instance, in addition to providing a greater number of overall cues, it provides the viewer with more realistic attributes or "handles" with which to store and consequently retrieve information. When color is used in this cueing role, its value could be predicted by the realism theories as well as by cue summation theory.

Based upon this vast accumulation of frequently inconsistent research data, a project was undertaken at the University of Pittsburgh to systematically investigate the relative effectiveness of the specific visual design factors of complexity and color realism, as they relate to certain pictorial memory tasks such as recognition, recall, and spatial analysis/synthesis. This paper is a summary of the findings and conclusions of those studies up until the present time.

RESEARCH MATERIALS

Much past research investigating the differences between color and monochrome visuals failed to take into account the fact that realistic color visuals contain intrinsically more information and consequently require more time for processing. In an attempt to resolve this methodological inconsistency as well as to more accurately assess the role of color in human information processing, Berry (1977) developed a set of stimulus materials for use in recognition studies. These consisted of 180 geographic scenery slides (120 stimulus slides and 60 distractor slides). The entire collection of materials was randomly divided into thirds. One third was retained as a realistic color group, a second third was recopied onto black and white slides and the remaining third was altered by photographic reversal to produce a non-realistic color group. By means of the photographic reversal process, the overall number of color cues could be held constant, while the degree of color realism could be manipulated.

A second set of materials was developed by Jesky (1984) for the purpose of testing hypotheses regarding recall memory. These materials were later modified by Berry (1990) to further test the realism hypotheses.The stimulus materials developed by Jesky were three sets of visuals, each produced in three visual formats; line drawing, black & white and color. Each set of visuals contained a different collection of common household items (32 per set) in a random arrangement. A nonrealistic version of these slides was produced by photographically reversing the color (realistic) slides of each set so that each color was systematically reversed to it's complementary value.

Additional modifications have been made to these materials over the course of this plan of research to accommodate or test specific, theoretical or related variables. These include digitized color reproductions, paired stimulus items and separation into specific object images.

The studies summarized in this paper used the research materials described above to test a variety of research hypotheses regarding visual complexity, color, cognitive styles and cerebral asymmetry as well as certain design factors and retention under either pictorial recognition or free recall situations. In the case of recognition memory experiments, the list learning procedure was employed, in which all subjects were first shown the set of 120 stimulus slides, individually for 500 msec. each. On a random basis, subjects were instructed to perform the masking task while viewing either the first 60 slides or the second 60 slides. The remaining, alternate 60 slides were viewed without performing the masking task. Subjects were subsequently presented a random distribution of all slides (stimulus and distractor) for 5 sec. each. During that time, subjects responded on a checklist either "old" (stimulus slide - seen before) or "new" (distractor slide - never seen). In those studies where pictorial recall memory was the task of interest, a free recall paradigm was employed. Subjects were presented with the stimulus materials for a brief study time (20 seconds) after which they were asked to recall as many items as they could from the stimulus images. The limited capacity of short term memory and rehearsal difficulty predicts that within a brief period of time (approximately four minutes), all items available in memory will have been recalled. Again, slight modifications were made in the methodology in some studies where variables of interest required specific presentation methods. In all studies, extreme care was exercised to maintain rigorous standardization of both materials and experimental procedures to ensure the generalizability of findings.

Over the time period of 1975 - 1990, a series of studies employed the materials and methods described to assess the effects of color and visual complexity. In all, fourteen studies have been completed or are nearing completion. A summary of the various studies conducted in this research program is presented in Table 1.

Insert Table 1 about here

SUMMARY of FINDINGS

This research project produced a number of significant findings and often confirmed the data of other studies. The results of each of these studies or groups of studies will be reviewed and summarized relative to the research variables of interest and the cognitive task under study.

By far, the largest group of studies addressed the role of color and specifically color realism as it related to a pictorial recognition task. These studies (Berry, 1977, 1982, 1983, 1984, 1990; Lertchalolarn, 1981; El Gazzar, 1984; Gobriyal, 1984; Waltz, 1990) utilized the recognition stimulus materials described above and focused on the comparative effects of different color presentation formats as one of their primary research questions. Each study compared three color modes, realistic color, nonrealistic color and monochrome (black & white) in the recognition task described. Other variables included in some of these studies are discussed elsewhere in this document. Early studies employed a simple analysis of correct recognition scores, however in most of these studies, analysis of the data was conducted via the method of Signal Detection Analysis. Signal Detection Theory (Swets, 1964) has been accepted as a reliable technique for assessing a subject's ability to describe the occurrence of discrete binary events and has been advocated as a means of reliably studying pictorial recognition memory (Snodgrass, Volvovitz, & Walfish, 1972; Loftus & Kallman, 1979; Loftus, Greene, & Smith, 1980; Berry, 1982).

The data obtained in this group of studies collectively indicated that significant differences generally exist between both realistic and nonrealistic color and black & white formats in favor of the color versions. This finding has consistently appeared with only two partial exceptions. The first is related to digitized pseudocolor, a variation of the nonrealistic color format that reduces the amount of color information systematically via a digitizing equation. In this case (ElGazzar, 1984, the realistic materials were superior to black & white, but the digitized images were inferior to the monochrome format. The second case (Berry, 1990; Waltz, 1990) incorporates the concept of cerebral laterality and suggests that nonrealistic color is superior to realistic color in right hemisphere processing only.

The generalized finding in these studies, however, lends strong support to the cue summation orientation, but does not necessarily support the notion that more realistic materials are intrinsically more effective in facilitating recognition.

The second largest group of studies focused on various cognitive style factors, specifically; field dependence, impulsivity/reflectivity and leveling sharpening. Generally, cognitive styles refer to "consistent modes of processing information" (Messick,1976), which are stable over time. Additionally, cognitive styles tend to incorporate a strong visual/perceptual component and have been significantly related to learning.

The most thoroughly researcher cognitive style is field dependence which was identified by Witkin, Oltman, Raskin & Karp (1971). This factor is functionally defined as the differential ability of individuals to separate figure from ground or to overcome "figural embeddedness". This ability of figural disembedding is generally considered to be representative of the more global ability to impose structure upon perceived information. Skill in this area has significant implications for the design of instructional materials.

The second cognitive style, impulsivity/reflectivity (Kagan, Rosman, Day, Albert & Phillips, 1964; Kagan, 1969), refers to the differential characteristic of some individuals to respond quickly yet inaccurately in a response situation (impulsives), while others respond more slowly but with more accuracy (reflectives). Research has indicated that this style has important consequences for instruction as well as the development of visual instructional materials because it also impacts upon an individual's ability to "read" a visual display.

The third cognitive style is leveling/sharpening which was identified by Holzman & Gardner (1960). Gardner & Long (1960) defined leveling/sharpening as "individual consistencies in memory organization as a function of assimilation between new stimuli and memories of stimuli experienced previously". In current practice, this construct refers to the ability of some learners to skim over significant differences or differentiations in content (levelers) while others tend to make discriminations which are not really valid (sharpeners).

Research which explored color realism and complexity in relation to cognitive styles was conducted by Wieckowski, (1980) (field dependence, impulsivity/reflectivity); Lertchalolarn, (1981) (field dependence); Gobriyal, (1984) (leveling sharpening) and Berry, (1984) (field dependence) as well as visual recall research by Jesky, (1984) (field dependence). With regard to studies which investigated field dependence it was found that in some cases field independent subjects benefitted more from color materials. In one instance, field independent individuals scored significantly higher on visual tasks generally, although in most cases, no significant relationship was found between field dependence and the use of color in stimulus materials. Other research which investigated impulsivity/reflectivity and leveling/sharpening did not find significant interactions with the color variable.

A third area of theory which was explored relative to color realism was hemispheric laterality. Studies by Berry, (1990) and Waltz, (1990) focused on the effect of variations in color information processing in both hemispheres of the brain, using different perceptual testing methods. Research on lateral cognitive functions has suggested that the two hemispheres of the brain are partially specialized for particular skills, with the left hemisphere oriented toward language and analytical skills and the right associated with visuospatial processing (Hellige, 1980; Sperry,1984; Gordon, 1986). Similarly, researchers (Davidoff, 1976; Jorgenson, Davis, Opella & Angerstein, 1981) have found evidence to conclude that color processing may be a right hemisphere activity.

The research conducted in this series of studies used two methods of localizing cognitive activity to different sides of the brain. Berry, (1990) used a verbal masking task to inhibit left hemisphere processing, while Waltz, (1990) used a perceptual method; projecting different images to the right and left visual hemifields. Results of both studies were highly similar and suggest the same conclusion. Significant differences existed for the processing of color information in the right hemisphere. Nonrealistic color seems to be processed primarily in the right hemisphere while other color and monochrome information is processed much more in the left hemisphere. Conclusions suggest nonrealistic color information is less verbally related than realistic color information and is thereby processed in a different manner and location.

A third study, Alfahad (1990), investigated visual recall memory with results which implied that color processing is localized more strongly to the right hemisphere. Color realism was not explored however.

Three studies investigated color realism or complexity relative to a visual recall task (Jesky, 1984; Alfahad, 1990; Berry, 1991). Memory research indicates that while recognition memory data describe the contents of memory, recall data provides insight into the methods whereby information is encoded into memory (Anderson & Bower, 1972). Since color cueing information is assumed to be a significant contributor to memory storage, it would appear important to investigate the effect of the type of color cues on the encoding process via a recall task. Two of the studies investigated the variable of visual complexity using color (realistic), black & white , and line drawing illustrations. In both cases, color was found to be superior to black & white, which was, in turn, superior to line drawing formats. The third study addressed the extended question of color realism with the same results as the first two studies regarding realistic color, monochrome and line drawings. Nonrealistic color was found to be inferior to all of the other treatments. It can be concluded from these findings that realistic color is reasonably effective in facilitating the encoding process, while nonrealistic color does not seem to play a useful role. To the contrary, nonrealistic color may actually inhibit the encoding process.

CONCLUSIONS

Based upon these findings, four general conclusions can be drawn:

1. All forms of color facilitate the recognition of visual material equally well. Both realistic and nonrealistic color materials are superior to monochrome materials in terms of their utility as cueing devices.
2. In recall memory tasks, realistic color cueing is most effective, followed by black & white and line drawing formats. Nonrealistic color cueing is the least effective method, and may inhibit encoding.
3. The processing of nonrealistic color information is lateralized to the right hemisphere while color and black & white processing is left hemisphere oriented.
4. Some evidence exists to suggest that field dependence may play a role in color information processing.

With regard to recognition memory, conclusion one supports cue summation theory, but nor necessarily the realism hypothesis. This may be explained by considering the nature of the recognition process. Recognition is a test of the contents of memory, not the effectiveness of the encoding/retrieval processes which store and access that information. In a recognition experiment, the subject does not need to search for the information, but rather match it with information already made available by the experimenter. In this case, the additional cues provide a larger pool of information with which to make the match. The relevance or realism of the cues is of little consequence, only the number are significant.

In recall memory, two cognitive processes are functioning; memory search and recognition. The subject must first search through memory to locate an appropriate image, then a recognition test must be performed. The nature of the search process reveals how the visual cues function. Information which is encoded using realistic cues will be compared with a previously stored, generic schema of the image and a reasonably complete match is made. Upon retrieval, the image will be easy to retrieve because it fits a common schema with few variations. In the case of nonrealistic cues,however, there may be a substantial mismatch between the actual appearance of the image and the generic schema previously stored in memory. Encoding will be more difficult because the new image differs significantly from that incorporated in the schema. This conflict may actually inhibit encoding and consequently make retrieval more difficult.

The data related to cognitive lateral functioning and color realism seem fairly clear, but conclusions drawn from them may be tenuous. It is apparent that nonrealistic color is processed primarily in the right hemisphere, possibly because it functions in a symbolic or coding role. In this way it could be thought of as simple images rather than more complex schemata. Realistic color or even black & white images, which are "familiar" have more complex semantic networks which include information and associations among images, verbal symbols and propositions. In short, the type of information which research has identified as being processed primarily in the left hemisphere (Gordon, 1986). Following this line of reasoning, it is apparent that nonrealistic, unfamiliar or unconventional images would have limited utility in left hemisphere processing and consequently are processed in a different manner in the right hemisphere.

REFERENCES

Anderson, J. R. & Bower, G. H., (1972). Recognition and retrieval processes in free recall. *Psychological Review. 79.* 97-123.

Alfahad, F. N. (1990). The interactive effects of cognitive functioning and visual realism on visual memory recall tasks. (Unpublished doctoral dissertation, University of Pittsburgh).

Berry, L.H. (1977). *The effects of color realism on pictorial recognition memory*. Paper presented at the annual convention of the Association for Educational Communications and Technology, Miami Beach, FL.

Berry, L.H. (1982). Signal detection analysis of color realism data. *Instructional Communications and Technology Research Reports. 13,*

Berry, L. H. (1983). *The effect of cross cultural variations and color realism on pictorial recognition memory*. Paper presented at the annual convention of the Association for Educational Communications and Technology, New Orleans, LA.

Berry, L. H. (1984). *The role of cognitive Style in processing color information: A signal detection analysis*. Paper presented at the annual convention of the Association for Educational Communications and Technology, Dallas, TX.

Berry, L. H., (1990). Effects of hemispheric laterality on color information processing. *Perceptual and Motor Skills. 71,* 987-993.

Broadbent, D.E. (1958). *Perception and Communication*. New York: Pergamon Press.

Broadbent, D.E. (1965). Information processing in the nervous system.*Science. 35,* 457-462.

Dwyer, F.M. (1967). Adapting visual illustrations for effective learning. *Harvard Educational Review,.37,* 250-263.

Dwyer, F.M. (1972). *A guide for improving visualized instruction..* State College, PA: Learning Services.

Dwyer, F.M. (1978). *Strategies for improving visual learning*. State College, PA: Learning Services.

Dwyer, F. M., (Ed.).(1987) *Enhancing Visualized Instruction - Recommendations for Practice* . State College: Learning Services

El Gazzar, A. L. I. (1984) A comparative study of the interactive effects of digitized and photographic image modes and color realism in a pictorial recognition task (Doctoral dissertation, the University of Pittsburgh). Pittsburgh, PA.

Gardner, R. W. & Long, R. I., (1960) Leveling-sharpening and serial learning. *Perceptual and Motor Skills. 10-* 179-185.

Gobriyal, S. Z. (1984). The effect of color mode and leveling-sharpening cognitive style on pictorial recognition memory. (Unpublished Doctoral Dissertation, University of Pittsburgh).

Gordon, H. W. (1986). The cognitive laterality battery: Tests of specialized cognitive functions. International Journal of Neuroscience, 29, 223-244.

Hellige, J. B. (1980). Cerebral hemisphere asymmetry: Methods, issues, and implications. ECTJ, 28, 83-98.

Holzman, P. S & Gardner, R. W., (1960). Leveling-sharpening and memory organization. *Journal of Abnormal and Social Psychology. 61*. 176-180.

Jesky, R. R.(1984). The interactive effects of pictorial presentation and cognitive style on a visual recall memory task (Doctoral dissertation, University of Pittsburgh, 1985). *Dissertation Abstracts International, 44*, 11077.

Kagan, J., (1969). *Matching Familiar Figures Test*. Cambridge: J. Kagen.

Kagan, J., Rosman, B. L., Day, D., Albert, J., & Phillips, W., (1964). Information processing in the child. *Psychological Monographs. 78*, 1.

Lertchalolarn, C. (1981). The interactive effects of color realism, pictorial attributes, and cognitive style on pictorial information processing. (Unpublished doctoral dissertation, University of Pittsburgh).

Messick, S., (1976). *Individualized Learning. New York: Jossey-Bass.*

Severin, W. (1967). Another look at cue summation. *AV Communication Review. 15*, 233-245.

Sperry, R. (1984). Consciousness, personal identity and the divided brain. Neuropsychologia, 22(6), 661-673.

Travers, R.M.G. (1964). The transmission of information to human receivers. *AV Communication Review. 12*, 373-385.

VanderMeer, A. W., (1952). Relative effectiveness of color and black and white in instructional films. Technical report # SDC 269-7-28. Port Washington, New York: Special Devices Center, Office of Naval Research.

Waltz, E. (1990). Cerebral laterality and color information processing. (Unpublished doctoral dissertation, University of Pittsburgh)

Wieckowski, T. (1980). *The interactive effects of color and cognitive style on a pictorial recognition memory task*. Paper presented at the annual convention of the Association for Educational Communications and Technology, Denver, CO

Table 1

Summary of Visual Complexity/Color Realism Studies

	VISUAL VARIABLE	LEARNING TASK	OTHER VARIABLES
Berry, 1977	Color Realism	Recognition	Delayed Retention
Wieckowski, 1980	Visual Complexity	Recognition	Field Dependence & Impulsivity/reflectivity
Lertchalolarn, 1981	Color Realism	Recognition	Field Dependence
Berry, 1982	Color Realism	Recognition	
Berry, 1983	Color Realism	Recognition	Multicultural
El Gazzar, 1984	Color Realism	Recognition	Digitized Pseudocolor
Berry, 1984	Color Realism	Recognition	Field Dependence
Jesky, 1984	Visual Complexity	Recall	Field Dependence
Gobriyal, 1984	Color Realism	Recognition	Leveling/Sharpening
Berry, 1990	Color Realism	Recognition	Cerebral Laterality
Alfahad, 1990	Visual Complexity	Recall	Cerebral Laterality
Waltz, 1990	Color Realism	Recognition	Cerebral Laterality
Berry, 1991	Color Realism	Recall	
Stanton, 1991	Color Realism	Recognition	Color Discrimination

Title:

Harnessing the Power of Interactivity for Instruction

Author:

Terry K. Borsook

Introduction

Since Halcyon Skinner patented what is believed to be the world's first teaching machine back in 1866 (Benjamin, 1988), a progression of technology has enticed visionaries and those simply looking for a better way to teach. A cavalcade of audiotapes, videotapes, television, filmstrips, slides, and computer-aided-instruction terminals has passed through the corridors of education, but have turned out to be disappointing. But the latest addition to our arsenal of educational media is very different. It is so different because it is not a medium such as paper or videotape; it can act, react and interact in ways none of its predecessors could–in ways that are, in fact, quite human-like.

What makes the computer unique is its potential for interactivity. Selnow (1988) sees computers as "interactants in one-to-one communication". Rather than begin with a definition of interactivity, let us begin by thinking about the experience of driving a car. Most people find driving a reasonably positive experience (at least when there is no traffic!). And, indeed, some find it a truly exhilarating experience (the interstate drivers who think they are on the Indianapolis Speedway track). But what makes driving, especially sportscars, fun? The answer, I propose, is interaction -- an extremely high degree of interaction. If it were simply speed and the experience of getting somewhere, a ride on the bus or subway should do the trick. But the fact is that these modes of transportation are not as fun because we cannot interact with them. We cannot communicate with them, nor them with us. Berlo (1960) claims that "Communication represents an attempt to couple two organisms, to bridge the gap between two individuals through the production and reception of messages which have meaning for both." The driver can send messages to the car, and the car can initiate and send messages back to the driver. The messages that are sent by the driver to the car depends on the messages sent by the car to the driver and vice-versa. Feedback determines the next response. That is interaction.

Consider how compelling videogames (and other computer games such as simulations) are. I believe that it is the tight, highly responsive environment of these games which gives them much of their appeal. Compare the experience of playing a videogame with merely watching someone else play one. Watching someone else play is not nearly as exciting as playing the game for oneself. Why? The missing ingredient is interactivity. And which do you think would be more effective, a video of a teacher explaining calculus or a personal tutor explaining it one-on-one? The point being made here is extremely important. It is the potential for interactivity that sets the computer apart from all other instructional devices. Text, color graphics, full-motion video, animation and sound can all be transmitted via media other than the computer. It is interactivity that makes the computer an educational tool like no other — with fantastic educational potential (see Ehrlich, 1990).

A primary purpose of this paper is to examine the concept of interactivity itself. The notion of interactivity is intuitively appealing and much has been written about how computers can be interactive but literature on the topic of interactivity itself is scant. This paper will explore interactivity in some depth. A second purpose is to see how the power of interactivity can be harnessed and put to work. A natural starting point for an inquiry into interactivity is with an examination of the richest and most versatile type of interaction of all: human-to-human communication.

Human-to-Human Communication:

To set a context within which to view human/computer interaction, it is useful to look at some models of human communication, the linchpin of which is interactivity.

The oldest and certainly one of the most popular and useful models of communication was actually born out of a need to describe electronic communication rather than human communication. The Shannon-Weaver model (Shannon and Weaver, 1949; Heinich, Molenda and Russel, 1989) has five essential components: 1) a source, 2) a transmitter, 3) a signal, 4) a receiver, and 5) a destination. If Bill wishes to ask Mary out for a date, Bill (the source) might call her up on the telephone (the transmitter) and ask her if she would go out with him (the signal) whereupon Mary's ears (the receiver) would receive the message which would then be transformed into something Mary's brain (the destination) could understand. It turns out that Mary has had a crush on Bill and so is delighted with the idea of going out with him. Her brain, previously the destination, now becomes the source and the process is approximately reversed.

The Shannon-Weaver model has proved helpful in looking at human/human communication but there are other models perhaps more tailored to the features of human/human communication than the Shannon-Weaver model which originated as a model of electronic communication.

One example is Berlo's (Berlo, 1960) model of communication. This model has six components: 1) the communication source, 2) the encoder, 3) the message, 4) the channel, 5) the decoder, and 6) the communication receiver.

Let us apply Bill and Mary's communication to this model. Bill is the Source. He, the source, has an intention to communicate. This intention is encoded by Bill's vocal apparatus. In more complex communication situations, the source and the encoder may be separate. For instance, the president of a corporation may be the source of some communication, but it is the lower managers which translate the intentions of the president and relay the message to others in the organization. Returning to Bill and Mary, Bill chooses to use his mouth to encode his intention (alternatively, he could have chosen to use his hand and written a letter) and also elects to use the telephone as a channel to transmit his message. Interaction between humans can be very rich.[1]

What Exactly Is Interaction?

Selnow's critical features of interaction:

Before proceeding, we must ask what is interaction, exactly? How will we know when we see it? Selnow (1988) offers a useful place to begin. He proposes the following three features as critical in characterizing interpersonal communication (p. 126):

1. Messages must be receiver specific. The source must tailor his or her communication for each receiver (*this implies, but goes beyond, recognition of the partner*).

---

[1] Human interaction here means where the parties are present in the same location Written personal communication would fit the definition of interaction except that it is delayed.

2. Message exchanges in the event must follow the pattern of a response contingent progression. Message exchanges are contingent on sequential feedback information (*this acknowledges communicative interdependence and flexibility, but also suggests activity and structure*).

3. The channel must provide for a two-way flow of information to accommodate the essential feedback component.

Berlo's Levels of Communicative Interdependence:

According to Berlo (1960), in communications, the source and the receiver are interdependent. That is, the source affects the receiver and the receiver affects the source. There are four levels of such interdependence:

1. *Definitional-Physical Interdependence.* At this level, the two parties depend on each other for their very existence. Clearly, without a receiver, there is no source, and vice versa. Often two parties are physically there and are talking to each other, but neither is responding to the other; they are each talking about very different things. In this case, they are physically interdependent–they depend on each other's existence in order to encode a communication. All too many instructional programs fall at this level. Software that permits only simple forward and backward movement resides at the definitional-physical interdependence level.

2. *Action-Reaction Interdependence.* The essence of this level of interdependence is the presence of feedback. According to Berlo, "Feedback provides the source with information concerning his success in accomplishing his objective. In doing this, it exerts control over future messages which the source encodes." (pp. 111-112). He uses the example of a speaker who tells a joke to an audience. If the audience laughs, he will probably continue telling jokes, if nobody even cracks a smile then he will either change his jokes or stop telling jokes altogether. Most educational software would probably fall into this category. The computer presents some information and can then branch to appropriate locations in the material depending on the student's actions.

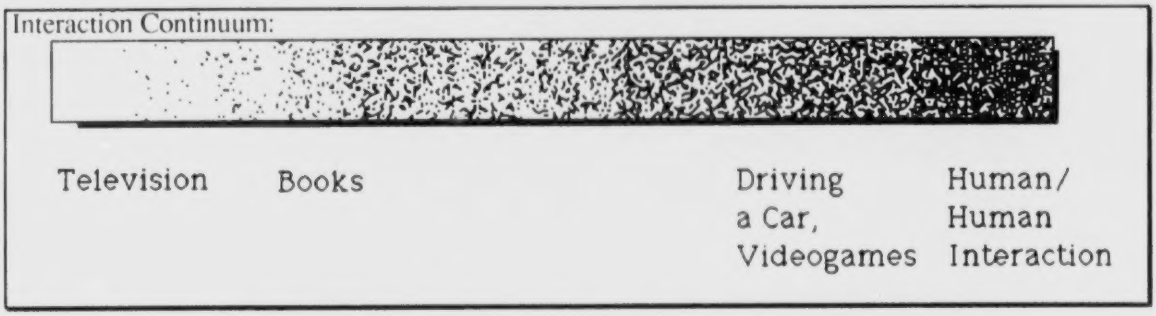
3. *Interdependence of Expectations*: *Empathy.* Empathy, in this sense involves making inferences of what a party is trying to communicate based on what that party does. It involves prediction.

4. *Interaction.* In Berlo's words, "If two individuals make inferences about their own roles and take the role of the other at the same time, and if their communication behavior depends on the reciprocal taking of roles, then they are communicating by interacting with each other." (p. 130). Berlo goes on to say, "When two people interact, they put themselves into each other's shoes, try to perceive the world as the other person perceives it, try to predict how the other will respond." Good computer simulations and games fit in this category.

Berlo's four levels of communicative interdependence hold many lessons to be learned by designers of instructional software. Consider a program that restricts interactivity to flipping through electronic pages sequentially. One may have plenty of opportunities to press the space bar to advance to the next page, but there is virtually no interactivity going on. One cannot give the computer much feedback. The computer does not exhibit any empathy – it has no way of predicting or adapting to your responses. Neither party can do much more than go back and forth.

Now, consider a program that allows the student to become a part of its world, where the computer seems to disappear and in its place is an entity who's own responses are highly related to the user's responses. An example of such a program is *SimCity*.[2]

It should be obvious by now that interactivity is not an on or off concept. In other words, we cannot say that some entity is either interactive or it is not interactive. Writers have attempted to classify communication into categories such as reactive (learners merely react to computer output), proactive (learners make all the decisions) and interactive (a give and take relationship) (Thompson and Jorgensen, 1989). The implication here is that once something is labelled as falling into one or the other category, it is assumed that it does not have any (or at least, many) attributes in common. The use of this taxonomy makes it difficult to recognize that just about every entity has, to some extent, a definable amount of interactivity. I believe a much more fruitful way (and one that lends itself to empirical studies) of looking at interactivity is to think of it as a property of just about every entity (here we restrict ourselves to common instructional entities such as books, video, etc.), where the amount of that property falls along a continuum from no interactivity to completely interactive. Some examples would be illuminating now.



Do you think books are interactive? They are. When reading a book, the words have some impact on the reader (at least that is what is supposed to happen!). The reader can respond by turning the pages, reading only certain passages, or going back to reread a difficult passage. The book is having an influence on the reader and the reader can issue responses consistent with that influence. What about television? Television can influence the viewer perhaps by changing attitudes towards a product, or by making one happy or one sad, but the viewer has no way of influencing the television in return, short of changing the channel. Perhaps the most interactive of all communications is that between humans in a face-to-face communication. In such situations, the messages are receiver specific, there is a rich feedback mechanism such as content of what is said, tone of voice, body language, etc., and the participants may adjust their messages in response to the others (Weller, 1988). For example, if, in conversing with Mary, Bill notices that his jokes are being met with either a blank stare or a frown, Bill can alter his messages as a result. So, Bill has influenced Mary who has in turn influenced Bill. This feedback loop is the hallmark of true interaction. We must remember, however, that people approach communications with others with various needs and goals. Bill and Mary may be talking to one another, but they are not responding to each other. That is, their goal is talking for the sake of talking. We perceive the illusion of interaction because they take turns talking. Critical ingredients such as the use of feedback and tailored messages are missing.

---

[2] SimCity is an absolutely incredible simulation program published by Broderbund software. It is a supreme example of how to successfully mix education and entertainment. Another excellent example of such a mix is Broderbund's Carmen Sandiego series of educational games.

Borsook's Recipe for Interactivity

What makes something interactive? While the term is freely bandied about, interactivity is an ill-defined concept, especially with respect to computers and its importance in instruction. There exists no cohesive definition of interactivity that can be used to guide the creation of computer-based instruction. This absence of guidelines and a dependence on loose, intuitive notions of what interactivity is and what it can do does the enterprise of instructional technology a disservice. The following is a list of variables that are believed to be the key ingredients of interactivity. Just as every instructional system may be said to possess a degree of interactivity that falls along a continuum from none to complete,[3] so each of these variables are present to some degree or another in each instructional system. In educational software one can find anywhere from none or a dash of a particular ingredient to any combination and quantities of ingredients. Interactivity increases to the extent that there is more or less of the following ingredients:

Immediacy of response. Think of the difference between conversing with a friend who is standing right before you and corresponding with her through the mail and you will appreciate the importance of immediacy of response to interactivity. In an educational program, if you wanted to know some more about, say a graph, you could click the mouse button and retrieve a text elaboration or another graph or a video, all immediately. In this sense, computers can allow for greater interactivity than humans! So, interactivity implies a sense of immediacy of response.

Non-sequential access of information. Interactivity allows for accessing and presenting information non-sequentially. Consider two different scenarios. Jack walks into a board room, makes his presentation from start to finish, accepts no questions, packs up, and leaves when finished. Tammy walks into another boardroom and begins her talk but tailors her presentation to the needs of the group. She fields questions, is able to restate some of her points in different ways and is responsive to the needs, concerns and anxieties of her audience. Which meeting would be more effective? Which meeting would you rather attend?

Adaptability. The third primary ingredient of interactivity, adaptability, is actually an implication of the non-sequential component. When two parties interact, there is adaptation taking place. We alter both what we say and how we say it dependent on to whom we are talking. We speak differently to a doctor, friend, professor, a new acquaintance, and a young child. More will be said about adaptability later on.

Feedback. The fourth ingredient of interactivity, feedback, is the information upon which change and adaptations are made. Interactivity stands on the shoulders of feedback. Feedback is what allows interactive systems to personalize and adapt instruction. Personalized feedback is the key to interaction (Selnow, 1988; Jonassen, 1985; Berlo, 1960). "Feedback provides the source with the information concerning his success in accomplishing his objective. In doing this, it exerts control over future messages which the source encodes" (Hooper and Hannafin, 1988; pp. 111-112; see also Schloss, Wisniewski and Cartwright, 1988). Smith (1972) points up the important distinction between feedback and knowledge of results.

---

[3] I propose that it is impossible for any system to be at either extreme of the continuum.

1. Knowledge of results does not facilitate performance if learners have incorrect hypotheses about what leads to correct performance. They may not realize they are attending to irrelevancies. They need to know the amount and direction of error as well as the outcome.

2. The information has to be incorporated into the pattern of control and to allow differentiation of relevant and irrelevant cues.

3. The learners always have some knowledge of results from their own sensory feedback, but it may not be sufficient.

So, the feedback package includes more than a response such as "right" or "wrong", it must also include directions for how to correct the situation if it is not right. Therefore, systems that are to be labelled highly interactive must, of necessity, also be said to be systems that utilize feedback extensively. To appreciate the difference feedback can make, think of how difficult it would be to drive a car blindfolded! To use an educational example, compare a computer program that presents material in the same way to every student regardless of how a particular student responds and which offers very few alternatives as to what could be done at any point with a computer program in which feedback to each student is highly responsive to the students' actions. If there is no feedback, there can be no interactivity.

Options. Built into the concepts of adaptability and feedback is the notion of options, our fifth component. Software could not be very adaptive if it did not offer a sufficient number of options to its user. Clearly, if a user can only press the space bar to sequentially page through screen after screen of information, the computer could not extract enough information that could be used to adapt instruction. One of the features that makes one-on-one human conversation so potentially interactive is the wide variety of options available to each conversant. In a conversation between two people that only permitted the use of only six words, there would not be much interactivity. Neither would there be much interactivity between a human and a computer program that offered such a restricted set of "words" with which to converse.

Bi-directional communication. Interactive systems implicate a channel that permits two-way communication. If a computer is doing most or all of the "talking" (i.e. presenting graphic after graphic or text after text), then interactivity is compromised. Just as we would not be satisfied with a telephone connection that only allowed for a one-way conversation, so we should not be satisfied with a computer program that restricts two-way interaction.

Interruptability. Related to *options*, is a party's freedom to initiate a course of action that is removed from previous interactions or to interrupt a currently running sequence. Imagine not being able to change a course of conversation with someone or interrupt a long-winded soliloquy. One would probably not feel a sense of high interactivity in such a case. Related to interruptability is the notion of *grain-size.* Grain-size refers to the length of time required of a given sequence before allowing further input. If a user cannot interrupt and/or initiate an action, then interactivity is low. A program that requires users to watch a fifteen-minute video before allowing responsive action is clearly not doing its part to introduce interactivity to educational systems.

For completeness, it would now be appropriate to turn to an examination of *learner control*, and *adaptive instruction*, two concepts that have been associated with enhanced interactivity. What we will see is that only adaptive instruction truly takes advantage of the computer's power by accounting for different individual needs and by lending a guiding hand through the labyrinthine paths of knowledge.

Learner Control

When we talk about computer programs that allowing for greater interaction between themselves and the learner, the implication is that those programs hand over some degree of control of the learning experience to the learner. The computer is not the only instructional device that permits the user to control the experience; books, for example, allow a great deal of latitude with which readers can obtain information. What makes the computer unique is its ability to respond to the learner. A responsive device gives learner control unprecedented powers. The question is, is this good?

The advantages of learner control include greater individualization of instruction, increased sense of personal responsibility for learning, and the potential to optimize learning efficiency (Hannafin and Colamaio, 1987). The empirical data, however, suggest that the story is not a simple one. What we will see is that total learner control is beneficial only to those who are already somewhat knowledgeable about a domain or who are generally high academic achievers. It will be argued that by increasing learner control, we are not necessarily increasing interactivity.

A study by Avner, Moore and Smith (1980), divided students into two groups: a computer-control condition where students had to answer questions and could not advance through the lesson if their answers were not correct, and a learner-control condtion where students could advance whenever they chose. There were no significant differences between the computer-control condition and the learner control condition when laboratory procedures involved simple instruction following. There was, however, a significant difference between the groups when the students had to make decisions that required an understanding of underlying principles. Those students who were forced to give a response (computer control), performed significantly *better* than those not forced to do so (learner control). Presumably the students were not knowledgeable about the lab principles before sitting down at the computer. They needed guidance and by providing them with the opportunity to skip over important questions, they had been given "enough rope to hang themselves with." This phenomenon could be called the *too-much-rope syndrome*.

Schloss, Wisniewski and Cartwright (1988) found that freedom to either review or not review material did not significantly affect performance on a posttest. The important factor was whether or not students received cumulative feedback on their performance. Learner control was more effective than program control of review presentation only when feedback was given. According to the researchers, the feedback provides "the learner with important information necessary in making informed choices."

These studies make a point which is becoming increasingly obvious: Learner control is good only for the informed learner. Whether that information comes in the form of performance feedback or in the form of previously learned knowledge does not appear to be as critical as whether information is available or not. Too much control over the branching of instruction may even lead learners to acquire negative attitudes towards the lesson (Gray, 1977).

The learner control literature seems to confirm the concept that it is the extent to which knowledge is applied to new learning which determines what and how much is learned (Jonassen, 1985b, 1988a, 1988b; Carrier and Jonassen, 1988; Ausubel, 1960). Students who navigate blindly through the myriad paths of instruction materials fall victim to the *too-much-rope syndrome* described earlier. Students who access previously acquired knowledge know what they need to know, know how to get it, and know what to do with it. Students with prior knowledge of a subject can perform equally well under learner as under computer control (Gay, 1986). The corollary is that students who do not bring previously acquired knowledge to a lesson need guidance. Computer-guided branching enhances interaction because without it, the computer either follows a predetermined path (low interaction), branches contingent on the learner's answers (moderate interaction), or turns over full control of branching to the learner (moderate interaction because the computer is not influencing [i.e. communicating] its recommendations) — it is simply reacting to the learner's input. Computer-guided branching maintains the spirit of true interaction which involves receiver specific and influenced messages, and a reciprocal give and take.

Adaptive Instruction

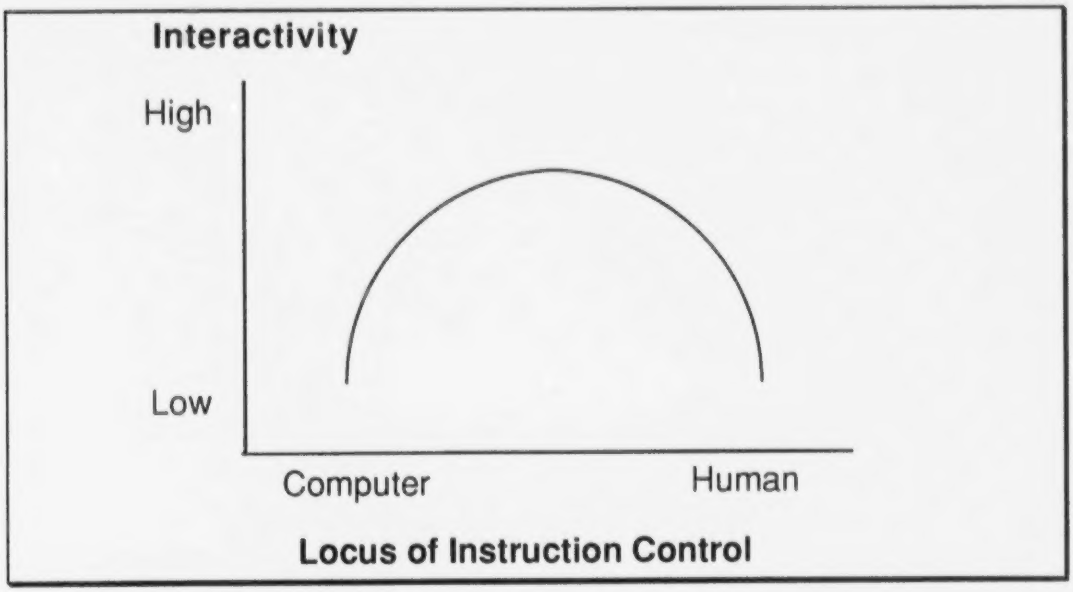
Now that we have decided that feedback and guided (but not forced) branching is advantageous over giving learners free rein, we are on the doorstep of another different, yet related field called *adaptive instruction*. Adaptive courseware is designed to dynamically alter presentation of instruction as a result of learner idiosyncracies. Elaborate and highly time-consuming models for the creation of adaptive courseware have been proposed (eg. Carrier and Jonassen, 1988), but the idea is simple. Let us consider the lowly, yet eminently useful book. The book cannot change because a student tends to be field-dependent rather than field-independent. It can neither assess the student's knowledge nor the student's personality. What is printed is printed. It cannot change. Now let us consider a good human instructor. She can note that the student knows about and likes baseball and can attempt to frame examples in baseball terms[4]. She can note that reading comprehension seems to be a problem and so she can provide remedial support in this area. The student provides feedback to her enabling her to fine-tune her responses and her responses, in turn, influence subsequent feedback. This feedback loop is one of the hallmarks of interaction. To the extent that we consider adaptation to enhance interaction, and to the extent that we consider enhanced interaction beneficial to the student, then we could confidently assert that adaptation is indeed beneficial to the student.

Contrary to claims and implications in the literature, increasing learner control is not commensurate with increasing the quality or quantity of interactivity offered by an instruction system. Research has shown that allowing all but the most bright and knowledgeable free rein at controlling sequencing, pacing, amount of practice and level of difficulty (Higginbotham-Wheat, 1990) results in disappointing performance. Even affective measures showed no advantage for such systems (Lahey, 1978). Why is this so?

---

[4] Of course, some subjects, such as Math and Statistics are more amenable to a baseball analogy than other subjects such as English or History, but the point is that the teacher can make that assessment.

Although the area of learner control is complex and there exists different kinds of learner control (Higginbotham-Wheat, 1990), the idea is quite simple: By giving learners greater control over various aspects of instruction such as pacing, and sequencing, they can tailor the instruction to their own style of learning, thereby enhancing the efficacy and efficiency of learning. Furthermore, this greater freedom should motivate students which should enhance learning even more. Unfortunately, all but the most bright, knowledgeable and/or motivated students simply get lost or frustrated — the too-much-rope syndrome. I argue that the reason for the disappointing results is that learner control simply shifts the locus of control from the computer (where the computer controls all aspects of the learning experience) to the learner. As locus of control shifts from one party to another, interactivity is diminished. Indeed, the very term *locus* of control implies an imbalance that is antithhesis to true interactivity.



Adaptive or guided instruction simulates a good, caring teacher who knows the material, and who takes the time and makes the effort to learn about and capitalize on each individual student's strengths, weaknesses and overall idiosyncracies. The student is encouraged, but not forced, to follow the teacher's suggested path. The student responds to the teacher and the teacher responds back to the student. Each response is contingent on the other's response, with personalized feedback the guiding light. This is interactivity. It has all the requisite ingredients of interactivity.

In addressing the issue of how motivation and curiosity can be stimulated so that students will go beyond inefficient browsing to goal-oriented and motivating learning, Kinzie and Berdel (1990) state that simply handing over all instructional control to the student is not necessarily good. Most students need to be guided. The argument for learner control is predicated on the idea that each student knows best what he or she needs to learn. The problem is that most students lack self-regulation skills. The trick is to somehow maintain the benefits of learner control yet somehow circumvent its limitations. These authors propose that in a hypermedia system, "learners can be prompted during practice to look up words, ask for alternative presentations (diagrams, spoken words, etc.), take notes or make drawings, or move about, depending on system capabilities." Some of the examples they give follows:

Hint	When to Provide It
"That word, 'mandible,' is a new one. If you don't know it yet, check the dictionary for a definition."	When a dictionary entry is available
"From here, you can find out about the bees' *waggle dance* or about the long distances bees will fly to find food."	When there are multiple links available to related information.
"You might want to diagram the social structure of a bee colony, to help you remember. Use the drawing tablet."	After user has been exposed to the basic information needed for such a diagram.

Using such a technique we have the best of both worlds. To the student it is like having a caring, patient teacher sitting right there. Learner control is still available. The sequence of exposure is not fixed as it is with a videotape — the student is free to explore. Yet the computer always has a hand extended when the footing gets a little unstable or when the material being presented requires some extra support. Advancements in artificial intelligence cognitive science will help to personalize the guidance to the needs of the individual.

In this sense, the computer can be even better than a personal human tutor. Even the most willing and able of human tutors could not instantly call up a color picture of a particular patient's heart, present an audio recording of its beating sounds, provide background information on the patient and on the patient's disease, summon definitions of difficult terms, and show an animated sequence highlighting the flow of blood through the organ, all the while monitoring the student's progress, making suggestions and remaining endlessly patient!

The point being made here is important enough to restate explicitly: *We should aim to emulate in computer instruction systems, to the greatest extent possible, the richness and flexibility found in human/human interaction while concomitantly exploiting the unique capabilities computers have to offer. That means making computer and student partners in the critical art and science of education.*

Bibliography

Allen, B.S., & Carter, C.D. (1988). Expert systems and interactive video tutorial: Separating strategies from subject matter. Journal of computer-based instruction, **15**(4), 123-130.

Ambron, S., & Hooper, K. (Eds.) (1988). Interactive Multimedia. Redmond, WA: Microsoft Press.

Axelrod, S. et al. (1987). Effects of video games as reinforcers for computerized addition performance. Journal of Special Education Technology, **9**(1), 1-8.

Barker, J. (1988). Interactive learning--freedom to think? Transition From Education Through Employment, Oct., 12-13. (Available from Personnel Pulications Ltd., 1 Hills Place, London, W1R-1AG, England.)

Berlo, D. (1960). The process of communication: An introduction to theory and practice. New York: Holt Rinehart and Winston.

Bloom, B.S. (1971). Mastery learning. In J.H. Block (Ed.), Mastery learning: Theory and practice. New York: Holt, Rinehart & Winston.

Carrier, C. (1984). Do learners make good choices? Instructional Innovator, **29**, 15-17.

Carrier, C., Davidson, G., & Williams, M. (1985). Selection of instructional options in a computer-based coordinate concept lesson. Educational Communications and Technology Journal, **33**, 199-212.

Carrier, C. A., & Jonassen, D.H. (1988). Adapting courseware to accomodate individual differences. In D.H. Jonassen (Ed.) Instructional designs for microcomputer courseware. Hillsdale, NJ: Lawrence Erlbaum Associates.

Davies, D. (1988). Computer-supported co-operative learning systems: Interactive group technologies and open learning. Programmed Learning and Educational Technology, **25**(3), 205-15.

Dede, C. (1986). A review and synthesis of recent research in intelligent computer-assisted instruction. Int. Journal of Man-Machine Studies, **24**, 329-353.

DeJoy, J.K., & Mills, H.H. (1989). Criteria for evaluating interactive instructional materials for adult self-directed learners. Educational Technology, Feb, 39-41.

Derry, S. (1984). Strategy training: An incidental learning model for CAI. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Desilets, B.J. (1989), Reading, thinking, and interactive fiction (instructional materials). English Journal, **78**(3), 75-77.

Duchastel, P. (1986). Intelligent computer assisted instruction: The nature of learner control. Journal of Educational Computing Research, **2**(3), 379-393.

Ehrlich, A. (1990). Stabilizing the term of art. CBT Directions, Oct., p. 32.

Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. Journal of Educational Psychology, **78**, 225-227.

Gay, G. et al. (1988, Oct.). Perceptions of control and use of control options in compuer-assisted video instruction. Techtrends, **33**(5), 31-33.

Gray, S.H. (1987). The effect of sequence control on computer assisted learning. Journal of Computer-Based Instruction. **14**(2), 54-56.

Hannafin, M.J., & Colamaio, M.E. (1987). The effects of variations in lesson control and practice on learning from interactive video. Educational Communications Technology Journal. **35**(4), 203-212.

Hazen, M. (1985). Instructional software design principles. Educational Technology, Nov., 18-23.

Heinich, R., Molenda, M., & Russell, J.D. (1989). Instructional media and the new technologies of instruction, 3ed, New York: Macmillan Publishing Company.

Huddleston, J.H.F. (1984). What individuality means for systems design. Behaviour and Information Technology, **3**(1), 85-91.

Jonassen, D.H. (1989). Hypertext/Hypermedia. Englewood Cliffs, NJ: Educational Technology Publications.

Jonassen, D.H. (1988). Integrating learning stategies into courseare to facilitate deeper processing. In D.H. Jonassen (Ed.) Instructional designs for microcomputer courseware. Hillsdale, NJ: Lawrence Erlbaum Associates.

Jacobson, E., & Thompson, M. (1975). Self-managed learning using CAI. Paper presented at the annual meeting of the American Educational Research Association. Washington, D.C.

Johansen, K., & Tennyson, R. (1983). Effects of adaptive advisement on perception in learner-controlled, computer-based instruction. Educational Communications and Technology Journal, **31**, 226-236.

Kern, G.M., & Matta, K.F. (1988). The influence of personality on self-paced instruction. Journal of Computer-Based Instruction. **15**(3), 104-108.

Kinzie, M.B., Sullivan, H.J., & Berdel, R.L. (1988). Learner control and achievement in science computer-assisted instruction. Journal of Educational Psychology. **80**(3), 299-303.

Lahey, G.F. (1978). Learner control of computer-based instruction: a comparison to guided instruction. Paper presented at the Annual Meeting of the Association for the Develoopment of Computer Based Instructional Systems. Dallas, TX.

Lahey, G.F., Crawford, A.M., & Hurlock, G. (1976). Learner control of lesson strategy: some tentative results. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Laurillard, D.M. (1984). Interactive video and the control of learning. Educational Technology, Jun, 7-15.

Lindner, R. W., & Rickards, J.P. (1985). Questions inserted in text: Issues and implications. In D.H. Jonassen (Ed.) The technology of text, vol II. Englewood Cliffs, NJ: Educational Technology Publications.

McCann, P. (1981). Learning strategies and computer-based instruction. Computers & Education, **5**, 133-140.

Matta, K.F., & Kern, G.M. (1989). A framework for research in computer-aided instruction: Challenges and opportunities. Computers and Education, **13**(1), 77-84.

McCann, P. (1981). Learning strategies and computer-based instruction. Computers & Education, **5**, 133-140.

Merrill, D. (1980). Learner control in computer based learning. Computers & Education, **4**, 77-95.

Merrill, D. (1984). What is learner control? (ERIC Document Reproduction Service No. ED 298 905).

Merrill, J. (1987). Levels of questioning and forms of feedback: Instructional factors in courseware design. Journal of Computer-Based Instruction, **14**(1), 18-22.

Midoro, V., et al. (1988). Interactive video and artificial intelligence: A convenient marriage. Programmed Learning and Educationl Technology, **25**(4), 299-309.

Morrison, G. R., Ross, S.M., O'Dell, J.K., & Schultz, C.W. (1988). Adapting test presentations to media attributes: Getting more out of less in CBI. Computers in Human Behavior, **4**, 65-75.

Niemiec, R., & Walberg, H.J. (1987). Comparative effects of computer-assisted instruction: A synthesis of reviews. Journal of Educational Computing Research, **3**(1), 19-37.

Reiser, R. (1984). Reducing student procrastination in a personalized system of instruction course. Educational Communications and Technology Journal, **32**, 41-49.

Ross, S.M. (1984). Matching the lesson to the student: Alternative adaptive designs for individualized learning systems. Journal of Computer-Based Instruction. **11**(2), 42-48.

Ross, S.M., McCormick, D., & Krisack, N. (1986). Adapting the thematic context of mathematical problems to student interests: Individual versus group-based strategies. Journal of Educational Research, **79**, 245-252.

Ross, S.M., & Morrison, G.R. (1988). Adapting instruction to learner performance and background variables. In D.H. Jonassen (Ed.) Instructional designs for microcomputer courseware. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ross, S.M., Morrison, G.R., & O'Dell, J.K. (1988). Text density level as a design variable in instructional displays. Educational Communication & Technology Journal, **36**(1), 103-115.

Ross, S., & Rakow, E. (1981). Learner control versus program control as adaptive strategies for selection of instructional support on math rules. Journal of Educational Psychology, 73, 745-753.

Sasscer, M. F., & Moore, D.M. (1984). A study of the relationship between learner-control patterns and course completion in computer-assisted instruction. Programmed Learning and Educational Technology, **21**(1), 28-33.

Scandura, J.M. (1984). Cognitive instructional psychology: System requirements and research methodology. Journal of Computer-Based Instruction, **11**(2), 32-41.

Schloss, P.J., Wisniewski, L.A., & Cartwright, G.P. (1988). The differential effect of learner control and feedback in college students' performance on CAI modules. Journal of Eduational Computing Research. **4**(2), 141-149.

Seidel, R., Wagner, H., Rosenblatt, R., Hillelsohn, M.J., & Stelzer, J. (1978). Learner control of instructional sequencing within an adaptive tutorial CAI environment. Instructional Science, **7**, 37-80.

Selnow, G. W. (1988). Using Interactive Computer to Commuicate scientific Information. American Behavioral Scientist, **32**(2), 124-35.

Shim, J.P. (1988) The visual interactive computer-aided approach to teaching zero-sum/nonzero-sum games. Social Science Computer Review, **6**(3), 392-397.

Steinberg, E. R. (1989). Cognition and learner control: A literature review, 1977-1988. Journal of Computer-Based Instruction, **16**(4), 117-121.

Steinberg, E.R. (1977). Review of student control in computer-assisted instruction. Journal of Computer-Based Instruction, **3**(3), 84-90.

Thompson, J.G., & Jorgensen, S. (1989). How interactive is instructional technology? Alternative models for looking at interactions between learners and media. Educational Technology, Feb, 24-26.

Underwood, J. (1989). Hypermedia: Where we are and where we aren't. CALICO Journal, **6**(4), 23-26.

Underwood, J. (1989). Hypercard and interactive video. CALICO Journal, **6**(3), 7-20.

Weller, H.G. (1988). Interactivity in microcomputer-based instruction: Its essential components and how it can be enhanced. Educational Technology, Feb., 23-27.

Wilcox, W.C., Richards, B.F., Merrill, M.D., Christensen, D., & Rosenvall, J. (1978). Learner control of the number of instances in a rule-using task. (ERIC Document Reproduction Service No. ED 201531). Sponsoring agency: National Science Foundation, Washington, D.C.

Title:

Doctoral Dissertation Research in Instructional Design and Technology from 1976 through 1988

Authors:

Edward P. Caffarella
Steven G. Sachs

Doctoral Dissertation Research in Instructional Design and Technology from 1976 through 1988

1991 AECT Annual Convention

Edward P. Caffarella
Professor
College of Education
University of Northern Colorado
Greeley, Colorado

Steven G. Sachs
Associate Dean
Instructional Technologies & Extended Learning
Northern Virginia Community College

Doctoral Dissertation Research in Instructional Design and Technology from 1976 through 1988

1991 AECT Annual Convention

Edward P. Caffarella
University of Northern Colorado
Steven G. Sachs
Northern Virginia Community College

This study is an analysis of the doctoral dissertation research completed by students in instructional design and technology (ID&T) programs from 1977 through 1988. This doctoral research was completed at forty-six institutions throughout the United States. The study includes several analyses of doctoral dissertation research. The first section is an analysis of the number of dissertations completed during each of the twelve years. The second section deals with the forty-six institutions and the number of doctoral degrees in ID&T awarded by each. The third section is a historical analysis of dissertation research in ID&T. The fourth section is a comparison of the research during the current period with dissertation research completed since 1921. The last section is a comparison of the ten most prolific ID&T programs over the past seventy years.

Research Design

The analysis is based largely upon the data presented in the recent Association for Educational Communications and Technology publication Doctoral Research in Instructional Design and Technology: A Directory of Dissertations, 1977 - 1986 (Caffarella and Sachs, 1988) and Doctoral Research in Instructional Design and Technology: A Directory of Dissertations, 1987 - 1988 (Caffarella, 1991). These two publications contain listings of dissertations completed by students in ID&T programs in the United States. A more complete analysis of the dissertation research is contained in two publications by Caffarella & Sachs (1990a, 1990b).

These publications and this presentation provide an opportunity to better define the research in the field of instructional design and technology thereby helping the profession to identify valuable studies and research trends. By building upon the research of others, future researchers can advance the field more collectively than could be done by one individual working in isolation.

The data on dissertations was supplied directly by the universities in the United States offering doctoral programs in instructional design and technology. These programs are offered under a wide variety of titles including instructional development, educational media, instructional technology, and instructional systems. The basic list of institutions was taken from the list of doctoral and masters programs in the Educational Media and Technology Yearbook (Miller and Mosley, 1986). These lists were checked against the membership list for the Professors of Instructional Design and Technology to make the list as comprehensive as possible.

Each participating institution supplied the following information for each doctoral student who graduated between 1977 and 1988.

Student's Name
Year of the Degree
Dissertation Title
Institution Name
Dissertation Chairperson

The intent was to make the listing inclusive rather than exclusive. The decision to participate and which dissertations to submit was left to the individual institutions. The listings include virtually all doctoral programs in instructional design and technology.

The list for dissertations completed from 1977 through 1986 was collected during the fall of 1987 and the winter of 1988. The data for 1987 and 1988 was collected during the spring of 1989.

Dissertations Completed

During the period 1977 through 1988 there were 1518 dissertations completed at forty-six institutions (see figure 1). From 1977 through 1988 the number of dissertations completed each year ranged from a low of 106 in 1980 to a high of 149 in 1983 and 1985.

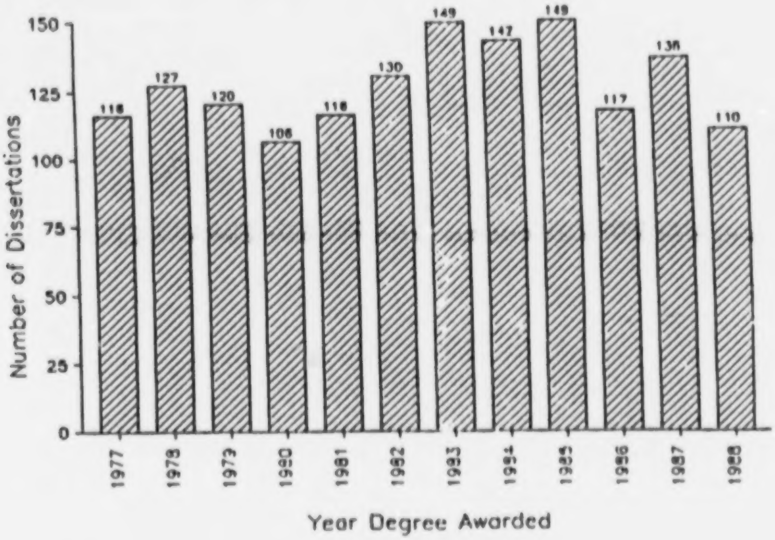


Figure 1. ID&T Doctoral Dissertations for Each Year, 1977-1988

Institutional productivity during the twelve year period showed that forty-nine percent of the graduates did their studies at only seven institutions (see table 1 & figure 2).

Table 1

Number of Dissertations for Each Institution

	Tot	Cum Tot	%	Cum %
Indiana University	216	216	14.23	14.23
Univ. of S. California	108	324	7.11	21.34
Boston University	107	431	7.05	28.39
Syracuse University	88	519	5.80	34.19
University of Pittsburgh	80	599	5.27	39.46
Florida State University	79	678	5.20	44.66
Brigham Young University	71	749	4.68	49.34
Others	769	1518	50.66	100

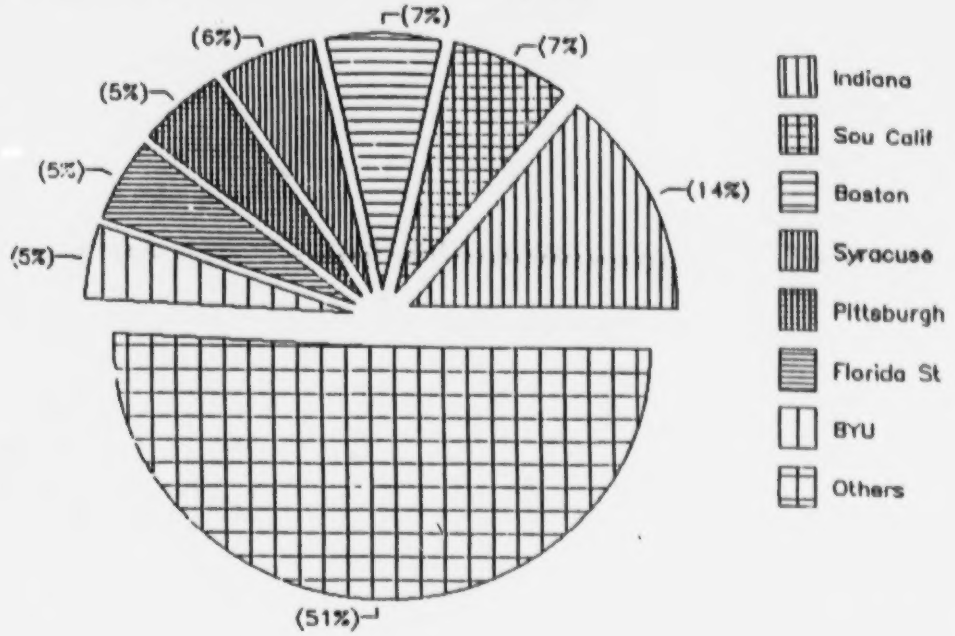


Figure 2. Seven Most Productive ID&T Institutions, 1977-1988

The largest number of graduates were from Indiana University with 216 dissertations. The other institutions in the six most prolific included the University of Southern California, Boston University, Syracuse University, the University of Pittsburgh, Florida State University, and Brigham Young University.

There was very little variability in the group of institutions producing the largest numbers of doctoral graduates during the twelve year period (see table 2). Over the period Indiana University was among the top three producers during every year and first for seven of those years. Boston University and the University of Southern California were in the top three for six years. Syracuse University was in this group three times from 1977 through 1988.

Table 2

Most Productive Institutions for Each Year

1977	Indiana (27)	Syracuse (16)	Sou. Calif. (9)
1978	Indiana (24)	Sou. Calif. (12)	East Texas (12)
1979	Indiana (24)	East Texas (9)	Boston (9)
1980	Indiana (22)	Sou. Calif. (9)	Penn. State (7)
1981	Boston (16)	Indiana (12)	Syracuse (9)
1982	Indiana (25)	Florida St. (13)	Boston (10)
1983	Florida St. (17)	Sou. Calif. (16)	Indiana & Pittsburgh (13)
1984	Sou. Calif. (18)	Indiana (14)	Boston (11)
1985	Indiana (18)	Boston (12)	B. Young & Mich. St. (10)
1986	Indiana & Boston (12)		Syracuse (8)
1987	Illinois (21)	Indiana (16)	Sou. Calif. (9)
1988	Illinois (10)	Indiana & Brigham Young (9)	

Within the institutions, the numbers of graduates varied widely from year to year. For example, Indiana University had 27 graduates in 1977 but only 9 during 1988. The number of graduates from the University of Southern California ranged from 18 in 1894 to 2 in 1986. Boston University's number of graduates ranged from 5 graduates in 1977 to 16 graduates in 1981.

Approximately twenty-five percent of the dissertations were completed under the direction of eleven professors. Half of the dissertations were directed by thirty-six chairs. As might be expected most of these professors are faculty members at the seven institutions producing the largest number of dissertations. There were 147 faculty members who chaired only one or two dissertations. Thus, a relatively small number of professors are directing the preponderance of the instructional design and technology dissertations.

Earlier Compilations

There are several earlier studies that compiled lists of dissertations in the field of instructional design and technology (Huang, 1979, 1980; Kirschner, Mapes, & Anderton, 1975; Moldstad, 1956, 1958, 1959, 1961). The earliest listed dissertation was completed at the University of Chicago in 1921. Between then and 1960 there were 729 completed dissertations with most of these completed since 1950. From 1960 until 1970 the doctoral students completed 468 dissertations. The data from 1971 through 1975 has not been compiled by anyone but is probably around 450 dissertations. The current study includes 1518 dissertations. As shown in figure 3 there have been at least

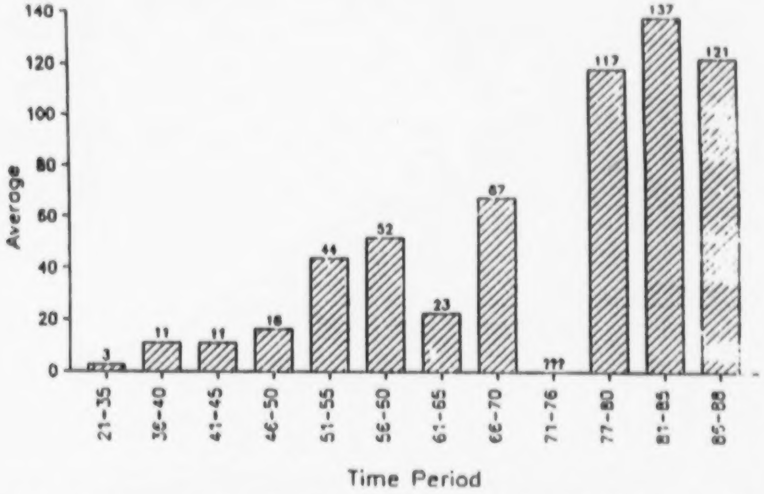


Figure 3. Average Number of ID&T Dissertations for Each Year during Each Period from 1921-1988

2695 dissertations completed in the field of instructional design and technology. If the 1971-1976 estimate is included this figure rises to approximately 3150. There has been a steady growth in number of dissertations completed each year.

A comparison was done of the most prolific institutions since 1921. The dissertations were divided into three groups including 1921-1956, 1960-1970, and 1977-1988 (see table 3).

Table 3

Ten Most Productive Institutions for Each Period

1921-1956	1960-1970	1977-1988
Columbia	Indiana	Indiana
New York	Syracuse	Sou. Calif.
Indiana	New York	Boston
Iowa	Sou. Calif.	Syracuse
Penn. St.	Michigan St.	Pittsburgh
Wisconsin	Ohio St.	Florida St.
Ohio St.	Columbia	Brigham Young
Chicago	Nebraska	East Texas
Stanford	Florida St.	Michigan St.
U. Michigan	Boston	Penn. St.

Comparing the top ten most prolific ID&T institutions for each period shows changes in the major institutions over time. Only Indiana University appears in all three lists. New York University and Ohio State University appear in the first two lists but not the last. Syracuse University, the University of Southern California, Michigan State University, Florida State

University, and Boston University appear on the last two lists but not the first.

The dissertation research produced by graduate students should be a major source for the field in building a knowledge base in instructional design and technology. Many times however the completed dissertations simply gather dusk on the shelves of university libraries. By making this research more readily available the field can capitalize on this excellent body of knowledge.

References

Caffarella, E. (in press) Doctoral research in instructional design & technology: a directory of dissertations, 1987-1988. Washington, D.C.: Association for Educational Communications and Technology.

Caffarella, E. P. & Sachs, S. G. (1988). Doctoral research in instructional design and technology: a directory of dissertations, 1977-1986. Washington, DC: Association for Educational Communications and Technology.

Caffarella, E. P. & Sachs, S. G. (1990a) "Doctoral Dissertations in Instructional Design and Technology, 1977 to 1988" Educational Technology: Research and Development, (38):3.

Caffarella, E. P. & Sachs, S. G. (1990b). "An analysis of doctoral dissertation productivity in instructional design and technology from 1977 to 1988." In B. Branyan-Broadbent & R. K. Wood (Eds.), Educational Media and Technology Yearbook 1990. Englewood, CO: Libraries Unlimited.

Huang, C. (1979). Monograph of doctoral studies in educational technology: 1960-1970. New York, NY: City University of New York, York College, Center for Educational Technology. (ERIC Document Reproduction Service No. ED 184 555)

Huang, C. (1980). Analysis of doctoral production in educational technology: 1960-1970. Educational Communications and Technology Journal, (28),1, 3-9.

Kirschner, C. D., Mapes, J. L. & Anderton, R. L. (1975). Doctoral research in educational media 1969-1972. Stanford, CA: ERIC Clearinghouse in Information Resources.

Miller, E. & Mosley, M. L. (1986). Educational media and technology yearbook 1986 (Vol. 12). Littleton, CO: Libraries Unlimited.

Moldstad, J. (1956). Doctoral dissertations in audio-visual education [1921-1956]. Audio-Visual Communication Review, 4(2), 291-333.

Moldstad, J. (1958). Doctoral dissertations in audio-visual education: supplement I [1956-1957]. Audio-Visual Communication Review, 6(1), 33-48.

Moldstad, J. (1959). Doctoral dissertations in audio-visual education: supplement II [1957-1958]. Audio-Visual Communication Review, 7(2), 142-153.

Moldstad, J. (1961). Doctoral dissertations in audio-visual education: supplement III [1958-1960]. Audio-Visual Communication Review, 9(4), 220-229.

Title:

The Facilitation of Vocabulary Acquisition by Embedding Explication in Expository Text

Author:

Loretta A. Cardinale

THE FACILITATION OF VOCABULARY ACQUISITION BY EMBEDDING EXPLICATION IN EXPOSITORY TEXT

The ability to read and understand text is vital for learning and remembering information. Learners typically have more difficulty reading expository than narrative text (Spiro & Taylor, 1980). A lack of motivation and interest may result in less time spent on reading and thus less experience with reading expository as well as narrative text (R. Anderson, Wilson, & Fielding, 1988).

Informative text that facilitates understanding, learning, and remembering uses certain text characteristics that facilitate learning from written materials. Standard characteristics of expository text include structure, coherence, unity, and audience appropriateness (Armbruster & T. Anderson, 1984). However, even texts that include those characteristics may lack features that enhance comprehension. Explication of information by familiar examples, causal relations, vocabulary usage, or analogy should enhance comprehension (Singer & Donlan, 1989). Specific types of explication may affect aspects of information processing differently. Those effects must be determined.

Information Processing and Learning from Expository Text

Mayer proposed four components of the encoding process: selection, acquisition, construction, and integration (Mayer, 1988; Mayer & Greeno, 1972; Weinstein & Mayer, 1986). It is each of these components that this study addressed by embedding specific types of explication: etymological, causal, and analogical.

Explication of Information

To explicate means to state directly, instead of requiring the reader to infer, organize, or construct relationships (Singer & Donlan, 1989). Learning will occur if the text directly explains the terminology or causal relations (Bean, Singer, Sorter, & Frazee, 1987). Thus, explication of information should be a concern for instructional designers.

Explication by Etymology in Text

Word etymologies are concerned with the origins and derivations of a word. Medical terminology is derived from about 4,000 Greek and Latin terms that form various prefixes, suffixes, and roots (Pepper, 1949). Familiarity with those Greek and Latin terms and an understanding of their use would save a student from memorizing more than 50,000 words and facilitate selective attention and acquisition.

Explication of Causal Relations in Text

Causality is an action that regularly produces an effect, and it means that there is a predictable or functional relationship between a cause and its effect (Schlick, 1949). Explicating causal relations should link one new term with another as it is encountered in the text. This linkage would organize the information and thus affect recall of simple arrangements and relationships (Mayer, 1988; Mayer & Bromage, 1980). Thus, in terms of Mayer and Greeno's (1972) model, explicating causal relations should enhance constructive processes. Retrieval processes may be enhanced as a result of linkages among knowledge from the text.

Explication by Analogy

The essence of analogical thinking is the transfer of knowledge from one situation to another by finding correspondence between the aspects of two different bodies of information (Gick & Holyoak, 1983). Where prior knowledge exists, an analogy can bridge the gap between the familiar and the unfamiliar by providing anchors for assimilation. Explication by analogy should enhance integration of new information with prior knowledge in long-term memory. This should improve both encoding and retrieval processes.

Momentum Effects from Sequencing

In most expository text, the author defines new terms before new factual information is presented. This is consistent with the role of prerequisite information in learning (R. Gagne, Briggs, & Wager, 1985). However, determining the effects of embedded explication requires a set of learning conditions that will favor the use of passages containing explications by the reader. There would be a point at which expository text is so easy that readers don't rely upon embedded explication or so difficult that readers will not attempt to learn the information. The arrangement of these experimental conditions must take momentum effects into account. Research investigating momentum effects has been associated with changes in the difficulty of instructional material that result from varied sequences (Rothkopf & Coatney, 1974).

Purpose

The purposes of this study were: 1) to examine experimental effects of presenting the section of the script containing explications of new terms before the new concepts (Vocabulary-First Sequenced Script) versus presenting that section after the new concepts (Vocabulary-Last Sequenced Script); 2) to examine the effects of etymological explications on acquisition of the meanings of new terms; 3) to examine the effects of explication of causal relations on construction and retrieval of information, and; 4) to examine the effects of analogical explication on integration of information in long-term memory.

Methods

To test the hypotheses, instructional materials were developed which presented new vocabulary before new facts (Vocabulary-First Sequenced Script) or presented new facts before new vocabulary (Vocabulary-Last Sequenced Script). For each sequence, one control group received filler material, and three groups received different types of explication resulting in a 2 (vocabulary sequence) X 4 (embedded explication) factorial design.

Subjects

A total of 184 undergraduate students at the Pennsylvania State University who were enrolled in elementary science education classes received 5 attendance points toward their final course grade for completing the materials and delayed test. All students completed a test delayed by 2 days. The eight treatments were randomly distributed among the students in each group.

Materials

The instructional content developed by Dwyer (1965) consisted of an 1,800 word instructional unit describing the parts of the human heart, the internal functions of those parts and the interrelated functioning of the heart. With the addition of explication passages written by the investigator, the instructional unit was expanded to 2,100 words.

Embedded Passages

Four types of embedded passages were included in the original text and were all located within the section discussing new terminology: The Heart and its Functions. This section was presented first to half of the subjects and last to the remaining half. Each embedded passage consisted of approximately the same number of words and was located at precisely the same point in the parallel texts. All explications referred to the same terms. The filler material, historical in nature, did not refer to the functions or parts of the heart.

Filler Material. Twelve passages consisting of historical material was added to the control group scripts in order to provide equivalent reading material with the same number of words. This filler material discussed the study of heart transplantation through the development of the artificial heart. For example:

> For the past twenty years or so doctors and surgeons have been studying the techniques of the transplanting of the human heart.

Explications. Each script contained one set of twelve explication passages that were etymological, causal, or analogical. They were placed at precisely the same points in all scripts. The Vocabulary-First Sequenced Script presented the section containing explicated vocabulary first while Vocabulary-Last Sequenced Script presented that section last.

Explication by Etymology. Etymological explications discussed the same terms as other explications. Information regarding the naming of the part of the heart or the meaning of prefixes and suffixes was stated within approximately the same number of words as the fillers and other explications. For example, the following explication discusses the meaning of prefixes and suffixes in pericardium:

> The word, PERICARDIUM can be broken down into two parts: PERI- and -CARDIUM. The root, CARDIUM comes from the Latin word for HEART while PERI comes from a Greek term meaning "around the outside".

Explicated Causal Relations. Explication of causal relations were approximately the same word count as the filler material and the other embedded passages. For example, the following statement links the function and characteristics of the pericardium to the term covered next in the script, the epicardium:

> The pericardium must be tough enough to protect the heart yet flexible enough to expand and contract. If it were damaged in some way, the epicardium would become shredded.

These statements were designed to show relationships between the names and functions of the parts of the heart. In this case, the epicardium is linked to the pericardium as it is in the heart.

Analogical Explication. The analogies were literal whenever possible, that is, they discussed both description and function. For example, both the appearance and function of the pericardium is included in this analogy:

> Plastic wrap for food, Saran Wrap or Handiwrap, is tough, transparent and elastic. It protects food. In a way, the pericardium is like plastic wrap.

The analogies were designed to activate prior, familiar knowledge and link it to the new terminology. The investigator attempted to generate meaningful, relevant analogies for the targeted subjects. Charts summarizing the script formats are found in Tables 1 and 2.

Delayed Test

A delayed test without an immediate test provides useful information about the use of embedded explication. The use of analogies has been shown to have effect on long-term memory (Gick & Holyoak, 1980; Simons, 1984). This supports the use of a delayed test without immediate testing. In addition, the interest of this study was to determine the amount of information held in long-term memory and not short-term memory. Immediate testing would tend to identify the amount of information insufficiently encoded to be stable in long-term memory (Ausubel, 1968). The goal of this study was not to determine the amount of information lost from short-term memory, but rather the information that had a degree of stability in long-term memory (i.e. after a delay period). The eighty-item test was structured into the following sections and given in the following order: Drawing, Identification, Terminology, and Comprehension. The types of questions in each of the four tests addressed a different aspect of the heart material. The Drawing Test required the student to draw a simple line picture of the heart and place the corresponding number of the 20 identified parts listed on the test. This tested the skill level of labeling or naming (R. Gagne et al., 1988). The Identification Test required a naming or labeling skill by recognition rather than cued recall. Without visual aids present during the reading, however, students had to infer the locations of structures from the information provided in the reading material. The Terminology Test, at the factual level, required the student to recognize the correct part of the heart, as an answer, for the function or description written in the question. The Comprehension Test, at the conceptual level, presented two functional and eighteen sequential questions that required the student to recall and apply information from different portions of the reading material. This test required knowledge of both the parts and functions of the heart along with information about circulation and the cycle of the heartbeat. The Kuder-Richardson 20 reliability coefficients are in Table 1.

Procedure

On the first day of the study, students were given 15 minutes to read the scripts. Two days after reading, students received the test packet. The drawing test was given for 5 minutes and the other three parts of the test were given for 10 minutes each. The four tests were given in the same order to minimize possible interference effects between tests. Students were instructed not to go back to the previous test.

Results

Overall, the effects of explication were significant on the drawing and terminology test. The effects of vocabulary sequence were only significant in the drawing test. The means and variances are in Table 2.

Cued Recall of Labels on the Drawing Test

On the drawing test, a list of 20 terms was provided to the student producing a cued-recall test condition. All terms were discussed in the script. The findings of the ANOVA of the group achievement on the drawing test resulted in significant differences among the treatment groups for both vocabulary sequence and embedded explication. There was no interaction between main effects. The summary of the ANOVA is in Table 3. From multiple comparisons, analogical-explication and causal-explication group means from the drawing test were significantly higher than the control group means.

Recognition of Labels on the Identification Test

The findings of the analysis of treatment group achievement on the identification test resulted in no significant difference among the score variances. Since subjects received no line drawings during the reading, this confirms the need for visual representation during instruction.

Recognition of Facts on the Terminology Test

The findings of the ANOVA of the group achievement on the terminology test resulted in significant differences among the treatment groups for explication. There was no interaction between main effects. The summary of the ANOVA is in Table 4. From the multiple comparison of the means, the groups receiving analogical explications and causal explications scored significantly higher than the control groups.

Recognition of Concepts on the Comprehension Test

Questions about the function and structure of the heart tested students at the conceptual level. The multiple choice test presented a high level of difficulty to students. In this study, reading material only once may have proved to be inadequate. The findings of the analysis of treatment group achievement on the comprehension test resulted in no significant difference among the score variances.

Discussion

The purpose of this study was to examine the effects of implicitly embedding short passages explicating new vocabulary on the acquisition of verbal information. Since the original text was conceptually dense and presented at least 12 new terms in 700 words, embedding explicative passages increased the ease of reading and enhance retrieval. Filler material had no effect other than to equilibrate reading time and momentum effects.

Overall, the treatment groups performed better on the measure of cued recall of labels, the drawing test, than the control groups that received filler material. The group mean scores for subjects receiving analogies and causal statements were significantly higher than the filler group mean scores on this test. Reading analogies and causal statements enhanced retrieval processes (Britton et al., 1989). These results indicate that richer propositions were encoded (J. Anderson & Reder, 1979).

The etymology groups did not score significantly higher than the control groups on any measures. This may be due to the subjects' lack of familiarity with etymologies (Richek, 1988). Training would have been required (Weinstein et al., 1988). The measure of cued recall depended upon retrieval processes. The etymologies may have induced students to attend and acquire the terms, but retrieval, which was measured, was not enhanced.

On the measure of cued recall, the students receiving the Vocabulary-Last Sequenced Script performed significantly better than the students receiving the Vocabulary-First Sequenced Script. This suggests that students who read the facts and concepts before the vocabulary relied on the embedded explications. This is consistent with Rothkopf's (1988) conclusions about momentum effects.

On the measure of factual verbal information, the terminology test, the groups receiving analogies scored significantly higher than all other groups. This supports the findings of other studies that examined analogical reasoning and indicates that analogies provide an opportunity for the learner to integrate new information with information held in long-term memory (Bean et al., 1985; Gentner, 1988; Gentner & Gentner, 1983; Gick & Holyoak, 1983; Hayes & Tierney, 1982; Simons, 1984; Vosniadou & Schlomm, 1988). The consistent performance of the groups receiving analogies on measures of cued recall and measures of recognition supports other research reporting the enhancement of both encoding and retrieval processes (Britton et al., 1989; E. Gagne, Yarbrough, Weidman & Bell, 1984).

Students reading the Causal Explication/Vocabulary-Last Sequenced Script treatment performed significantly better than students reading the etymologies on the measure of cued recall that tested naming. This supports the findings of previous studies that examined the effects of explicating causal relations (Bean et al., 1987; Herman et al., 1987; Pedersen et al., 1988). However, the Causal/Vocabulary-First students did not score significantly higher than the etymology groups. This supports Rothkopf's (1988) theory of momentum effects. The groups receiving the Causal Explication/Vocabulary-First Sequenced Script may have attended less to the explications while the groups receovomg the Causal Explication/Vocabulary-Last Sequenced Script may have attended to the explications and relied more on those reading aids.

Conclusions

The results of this study may be considered for the design of text-based instruction. To be effective, text-based explication must target mental processes. When construction is required by a task, then causal explications would benefit the learner by facilitating linkage between vocabulary. For example, learning the terminology for parts of a machine where each component causes an action and reaction would be facilitated by causal explication. Analogical explication would enhance the integration of new information with prior knowledge in situations where the vocabulary is unfamiliar, complex and abstract such as those found in microcomputer education. Particularly in cases where the information is conceptually dense, embedded explication can both relieve density and enhance processing (Singer & Donlan, 1989). The type of explication depends upon the goal of instruction, the method of testing, and the familiarity of the information (E. Gagne et al., 1984).

The results of this study suggest that training is necessary for the use of etymologies for vocabulary acquisition. Students must fully understand the structure of words, in particular, the role of prefixes, suffixes, and roots. They must also have some schema for foreign languages. A familiarity with classical languages may enhance vocabulary acquisition. This is an area for further research.

Recommendations for Future Research

Future research about learning from text requires specific measurement techniques that target specific aspects of information processing. Each of the treatments in this study could be replicated by modifying the tests, content, subjects, and conceptual density in order to determine specific processes that occur while learning.

Another area for future research would be to study the effect of training students to use implicitly embedded learning strategies. Biology instruction, with its extensive use of language, would be a valid area to attempt training in the use of etymologies. Perhaps if training could demonstrate the efficiency and application of knowing a few thousand prefixes and suffixes instead of tens of thousands of difficult terms, results would be significant.

Table 1. Kuder-Richardson Test Reliability Coefficients

Test	K-R 20	Item Type
Drawing Test (n = 20)	0.83	Cued Recall
Identification Test (n = 20)	0.81	Forced Choice
Terminology Test (n = 20)	0.83	Forced Choice
Comprehension Test (n = 20)	0.77	Forced Choice
Total-Criterion (n = 80)	0.92	Combination of above items

Table 2. Means and Variances of the Delayed Criterion Tests

Vocabulary-First Sequenced Script

Treatment Group

Test	Filler	Etymology	Causal	Analogy
Labels: Cued Recall				
Mean	3.00	3.52	4.61	4.70
Variance	6.00	7.62	9.18	4.93
Labels: Recognition				
Mean	7.09	7.70	7.65	8.74
Variance	12.96	6.60	14.59	7.02
Facts: Recognition				
Mean	7.43	7.65	8.35	9.04
Variance	5.43	12.67	10.43	14.14
Concepts: Recognition				
Mean	7.09	7.17	7.78	7.04
Variance	8.18	9.86	9.80	13.10

Vocabulary-Last Sequenced Script

Treatment Group

Test	Filler	Etymology	Causal	Analogy
Labels: Cued Recall				
Mean	3.78	4.13	5.48	5.56
Variance	7.45	9.92	6.97	5.43
Labels: Recognition				
Mean	7.34	8.39	8.39	8.78
Variance	6.76	12.53	3.80	14.74
Facts: Recognition				
Mean	7.30	8.43	8.78	9.43
Variance	8.47	8.88	9.73	15.44
Concepts: Recognition				
Mean	6.83	8.39	8.56	7.83
Variance	10.24	15.44	9.18	7.78

Table 3. ANOVA Summary Table for the Effects of Sequence and Explication on the Drawing Test.

Source	df	SS	F
Sequence Group	1	29.96	4.03*
Explication Group	3	106.10	4.92**
Sequence by Explication Interaction	3	0.36	0.02
Error	176	1266.00	
Total	183	1401.43	

* Significant in terms of the 0.05 level

** Significant in terms of the 0.01 level

Table 4. ANOVA Summary Table for the Effects of Sequence and Explication on the Terminology Test.

Source	df	SS	F
Sequence Group	1	6.3	0.59
Explication Group	3	86.7	2.70*
Sequence by Explication Interaction	3	4.9	0.15
Error	176	1875.1	
Total	183	1973.0	

* Significant in terms of the 0.05 level

** Significant in terms of the 0.01 level

REFERENCES

Anderson, R. C., Wilson, P., & Fielding, L. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly, 23*, 285-303.

Armbruster, B. and Anderson, T. H. (1984). *Producing a "Considerate" expository text: Or easy reading is damned hard writing.* Reading Education Report No. 46. Urbana: University of Illinois, Center for the Study of Reading.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York: Holt, Rinehart, & Winston.

Bean, T.W., Singer, H., and Cowan, S. (1985). Analogical study guides: Improving comprehension in science. *Journal of Reading, 29,* 246-250.

Bean, T., Singer, H., Sorter, J., and Frazee, C. (1987). Acquisition of hierarchically organized knowledge and prediction of events in world history. *Reading Research and Instruction; 26,* 99-114.

Britton, B. K., Van Dusen, L., Gulgoz, S., Glynn, S. (1989). Instructional texts rewritten by five expert teams: Revisions and retention improvements. *Journal of Educational Psychology, 81,* 226-239.

Dwyer, F.M. (1978). *Strategies for improving visual learning: A handbook for the effective selection, design and uses of visual learning materials.* State College, PA: Learning Services.

Dwyer, F.M. (1965). *An experimental evaluation of the relative effectiveness of selected visual illustrations in teaching science concepts to college freshman.* Unpublished doctoral dissertation, The Pennsylvania State University, State College, Pennsylvania.

Gagne, E., Yarbrough, D., Weidmann, C., & Bell, M. (1984). The effects of text familiarity and cohesion on retrieval of information learned from text. *Journal of Experimental Education, 52,* 207-213.

Gagne, R., Briggs, L.J., & Wager, W.W. (1988). *Principles of instructional design.* New York: Holt, Rinehart and Winston, Inc.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59,* 47-59.

Gentner, D. and Gentner, D.R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. Stevens (Eds.) *Mental Models*, Hillsdale, NJ: Erlbaum.

Gick, M. L. and Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12,* 306-355.

Gick, M. L. and Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Hawthorne, P. (1968). The Transplanted Heart, New York: Rand McNally & Company.

Hayes, D.A. and Tierney, R.J. (1982). Developing reader's knowledge through analogy. Reading Research Quarterly, 2, 256-280.

Herman, P.A., Anderson, R.C., & Pearson, P. D. (1987). Incidental acquisition of word meaning from expositions with varied text structures. Reading Research Quarterly, 22, 263-284.

Mayer, R.E. (1988). Learning strategies: An overview. In C.E. Weinstein, E.T. Goetz, & P.A. Alexander (Eds.), Learning and Study Strategies: Issues in assessment, instruction, and evaluation. New York: Academic Press, Inc.

Mayer, R.E. & Greeno, J. (1972). Structural differences between learning outcomes produced by different instructional methods. Journal of Educational Psychology, 64, 165-173.

Pepper, O. H. P. (1949). Medical Etymology, Philadelphia: W. B. Saunders Company.

Richek, M. A. (1988). Relating vocabulary learning to world knowledge. Journal of Reading, 32, 262-267.

Rothkopf, E. (1988). Perspectives on study skills training in a realistic instructional economy. In C. E. Weinstein, E. T. Goetz, and P. A. Alexander (Eds.) Learning and study strategies: Issues in assessment, instruction, and evaluation. New York: Academic Press, Inc.

Rothkopf, E. & Coatney, R. (1974). Effects of readability of context passages on subsequent inspection rate. Journal of Applied Psychology, 59, 679-682.

Schlick, M. (1949). Causality in everyday life and recent sciences. In H. Feigl & V. Sellars (Eds.), Readings in philosophical analysis. New York: Appleton-Century-Crofts.

Simons, P.R.J. (1984). Instructing with analogies. Journal of Educational Psychology, 76, 513-527.

Singer, H. & Donlan, D. (1989). Reading and learning from text. Hillsdale, NJ. Lawrence Erlbaum.

Vosniadou, S. & Schlomm, M. (1988). Explanatory analogies can help children acquire information from expository text. Journal of Educational Psychology, 80, 524-537.

Weinstein, C.E. & Mayer, R.E. (1986). The teaching of learning strategies. In M.C. Wittrock (Ed.) Handbook of research on teaching. Hillsdale, NJ: Lawrence Erlbaum.

Weinstein, C.E., Zimmerman, S.A. & Palmer, D.R. (1988). Assessing learning strategies: The design and development of the LASSI. In C.E. Weinstein, E.T. Goetz, and P.A. Alexander (Eds.) Learning and study strategies: Issues in Assessment, instruction, and evaluation. New York: Academic Press, Incorporated.

Title:

The Effects of Program Control, Learner Control, and Learner Control with Advisement Lesson Control Strategies on Anxiety and Learning from Computer-assisted Instruction

Author:

Randall P. Coorough

THE EFFECTS OF PROGRAM CONTROL, LEARNER CONTROL AND LEARNER CONTROL WITH ADVISEMENT LESSON CONTROL STRATEGIES ON ANXIETY AND LEARNING FROM COMPUTER-ASSISTED INSTRUCTION

Abstract

The purpose of this study was to examine the effects of three CAI locus of instructional control strategies (learner control, learner control with advisement and program control) on posttest performance and anxiety. The instructional content was a lesson addressing the effects of alcohol on the body. To examine the effects of the CAI lessons on learner anxiety and achievement, split-plot and pre- and posttest experimental designs were used. Subjects were undergraduate students randomly assigned to one of the three CAI treatment groups. All subjects completed the computer administered pre-assessment, CAI tutorial, and post-assessment. Pre- and post-assessment measurements consisted of a 24-item pretest and posttest along with a 10-item State-Trait Personality Inventory S-Anxiety subscale. Anxiety was assessed before, during, and after CAI. The major findings indicated that there were (1) no significant achievement differences between lesson control treatments; (2) significant decreases in subjects' state anxiety levels across measurement occasions that were the same across treatments; and (3) the learner control with advisement treatment required a significantly greater amount of time to complete the CAI lesson than was required by the learner control treatment.

Introduction

The microcomputer has evolved into an important tool for delivering instruction to students. An often cited advantage of computer-assisted instruction (CAI) is the control that can be given to individual learners over the lesson presentation (Hannafin,1984). Students using CAI may control the sequence, pace, depth of study, feedback, and presentation style of instruction thereby greatly accommodating individual learner requirements. Instructional control of CAI courseware has generally taken two forms: (a) externally controlled (learner control) instruction where learners have full or partial control over CAI lesson presentation (e.g. lesson sequence, pace or feedback); and (b) internally controlled (program control) instruction where CAI is preprogrammed by the designers leaving no control to learners (Hannafin, 1984).

Intuitively, many researchers have hypothesized that providing learners with greater control over instruction would result in greater individualization of instruction, higher achievement and more positive learner attitudes toward instruction. The research, however, in general, shows that learners exposed to learner control do not achieve more, effectively utilize control options, or manage their time more efficiently when compared to students using program controlled CAI (Carrier, 1984; Clark, 1984; Merrill, 1980; Steinberg, 1977, Tennyson, 1980). A more recent review of the learner control research suggests that the empirical evidence provides only mixed support for the benefits of learner control (Milheim & Azbell, 1988).

Several researchers have suggested that achievement from learner controlled CAI is significantly enhanced when some form of advisement or "coaching" is provided to learners (Hannafin, 1984; Ross, 1984; Tennyson & Buttrey, 1980, Tennyson, 1981). This variation of learner controlled CAI, according to Hannafin (1984), provides learners with advisement based upon the learner's performance, while maintaining learner control over instructional variables in the lesson. Advisement may include recommendations for the number of practice items to complete, content to be reviewed, or sequence to follow. Milheim and Azbell (1988) in reviewing the learner control with advisement literature found that groups receiving advisement, in general (1) had higher posttest means; (2) had more students reach mastery; (3) had longer time on task; (4) needed less instructional time; and, (5) needed fewer instructional instances.

Given the microcomputer's ability to provide individualized instruction, it is important to understand the interaction between CAI design strategies and learners. Many learner attributes have been implicated as playing a positive or negative role in learning via CAI. Previous research has indicated that students' prior achievement (Ross & Rakow, 1980), locus of control (Louie, Luick, & Louie, 1985), motivation (Seymour, Sullivan, Story & Mosley, 1987), and sex (Fennema & Sherman, 1977) are all contributing factors to a student's performance with CAI. The research also indicates that anxiety, a very critical learner characteristic, can have profound implications for learning during CAI (Sieber, O'Neil & Tobias, 1977).

Tobias (1979) describes anxiety as an affective state that may cause interference in cognitive processing during any of the three information-processing components: input, processing, and output. Anxiety reactions to a particular situation, called state anxiety, are characterized by subjective feelings of apprehension and tension, accompanied by arousal of the nervous system (Spielberger, 1966). A number of studies investigating the interactions between learner anxiety and CAI have indicated that certain levels of anxiety can be debilitating toward performance in CAI (Leherissey, O'Neil, & Hansen, 1971; O'Neil, 1972; O'Neil, Spielberger, & Hansen, 1969).

While anxiety can affect learning in CAI, Clark (1984) has reported that efforts to alleviate anxiety which inhibits performance from computer-based instruction have largely failed. Clark further suggests that it may be more effective to circumvent debilitating anxiety rather than attempting to curing it. Therefore, elements in CAI designs that may reduce anxiety would have both positive affective and cognitive implications (Clark, 1984). Tobias (1979) recommends instructional designs that provide organized lesson structure, small instructional steps, continual feedback, examples, and sufficient instructional review time.

While the effects of anxiety on CAI performance have been documented empirically, the interaction of CAI lesson control strategies, anxiety and resulting learner performance remains less clear however. It has been suggested that the more control a student has over a learning task, the better will be the performance and as control over learning material increases anxiety levels would decrease (Sieber et al., 1977). A number of researchers have attempted to determine the effect of CAI on performance and anxiety when learners were given control of various instructional variables (e.g. memory support, presentation of visual stimuli, feedback, and sequence of instruction) (Collier, Poyner, O'Neil, & Judd, 1973; Gallagher, O'Neil & Dick, 1971; Hansen, 1972; Judd, Daubek, & O'Neil, 1975). These studies generally suggested that providing learners with control of the instructional variables did not necessarily facilitate performance over imposed control groups, however, learner control to some extent, seemed to reduce a student's anxiety level.

While learner control in the above studies did not facilitate increased performance, it did affect anxiety. The purpose of this investigation was to examine the effects of program controlled, learner controlled and learner controlled with advisement locus of instructional control strategies on learning and state anxiety levels. Specifically, this study sought to determine whether a learner controlled CAI lesson with advisement would differentially affect anxiety and performance when compared to other CAI lesson designs.

Methods

Subjects

The study subjects were 106 undergraduate students from elective humanities classes at the University of Florida. The subjects in these classes represented a general cross-section of the University of Florida undergraduate population. Participation in the study was voluntary and subjects were given extra credit for participation.

CAI Lesson Content

The computer-assisted instructional task selected for this study was an instructional program that introduced information and assisted students in developing an awareness of alcohol and its psychological and physiological effects on the body. The CAI program was divided into four sections: (a) Alcohol the Drug; (b) Effects of Alcohol; (c) Metabolism of Alcohol; and (e) Alcohol Withdrawal. Each lesson section consisted of from eight to ten instructional frames and embedded questions. The majority of instructional frames also consisted of multiple layers of textual information that were progressively presented as students proceeded. Each of the three CAI treatments consisted of the same basic CAI tutorial program content with modifications to the locus of instructional control as discussed below.

Instructional Treatments

Three CAI lesson versions representing different CAI locus of instructional control design strategies were developed for the Macintosh computer. The three lesson control strategies included program control, learner control, and learner control with advisement.

Program Controlled CAI. The program control version presented the lesson in a predetermined sequence. Students proceeded through the lesson answering embedded questions and receiving response sensitive feedback after each answer was given. Instruction continued with correct answers while incorrect responses resulted in branching to a review of the content related to the question. Learners were given a second opportunity to answer the question correctly and then the lesson continued.

Learner Controlled CAI. The learner controlled CAI version provided students with control over lesson sequence, content review, number of practice questions, and the decision when to terminate the CAI lesson. Learner control subjects could make control decisions at points where the program control treatment directed the learning task. Students controlled the lesson sequence by choosing the order of sections studied and were able to move forward or backward in each section as well. Students answered embedded questions or skipped them proceeding to the next instructional frame. Formative feedback was provided for correct answers. Students giving incorrect responses were given the option of branching to and reviewing content or to proceed to the next section. When students chose to review a segment, the option to repeat the question a second time was given. If the student chose not to repeat the question, the lesson continued to the next section. After the second presentation of the question, the lesson continued.

Learner Control with Advisement CAI. The learner control with advisement group was presented with the same CAI lesson version as the learner control group except that advisement was provided at each decision point in the lesson. At the decision points, students in the learner control with advisement condition were presented with advisement as to the recommended lesson sequence, when to review content, and when to repeat questions. If students attempted to proceed through the lesson in a sequence deemed inappropriate by the designer, advisement appeared suggesting the most appropriate lesson sequence. In addition, when learners answered questions incorrectly, advisement appeared suggesting that content be reviewed. If students attempted to leave the lesson before finishing all sections they were advised to complete all content areas before terminating instruction.

Procedures

State Anxiety Scale assessments, pretest measurements, CAI lessons, and posttest measurements were conducted in a lab where subjects were seated individually at a Macintosh II workstation. Each Macintosh II was configured with a color monitor, keyboard, mouse and networked to a printer. There were a total of 16 Macintosh II workstations in the lab. Up to 16 subjects could be accommodated during a CAI session. While students were randomly assigned to treatment groups, they were allowed to sign up for any lab time scheduled, therefore, labs consisted of subjects from different treatment groups. Before starting the CAI lesson, subjects were given an introduction to the Macintosh Lab and then instructed how to use the mouse thereby enabling them to proceed through the HyperCard designed CAI lesson.

State Anxiety Scale Assessments. State anxiety was periodically assessed using the S-Anxiety subscale of the State-Trait Personality Inventory (STPI) developed by Spielberger (1979). The S-Anxiety subscale of the STPI, consisting of ten items, was administered three times via the Macintosh computer programmed with HyperCard. The session began with an initial assessment of state anxiety level (SA1) prior to beginning the CAI lesson and the pretest. Students responded to the ten questions by pointing with the mouse to the appropriate response (four possible responses) and then clicking the mouse button to select that response. Upon completing the initial state anxiety measurement, subjects were then branched to the pretest assessment. Subjects were later readministered the STPI state anxiety scale midway through the CAI lesson (SA2) and then again following administration of the posttest (SA3).

Pretest Procedures. Pretest measurements were conducted after the initial state anxiety assessment. The pretest consisted of 24 randomly ordered multiple-choice questions. The pretest was conducted via the Macintosh computer programmed by HyperCard. After the pretest, subjects were branched directly to the beginning of the CAI lesson.

Posttest Procedures. The posttest measurement, consisting of the same 24 pretest questions arranged in a different order, was conducted immediately following the CAI lesson prior to the administration of the final state anxiety inventory. The posttest was administered like the pretest.

Design and Data Analysis

In this study a split-plot experimental design with three between-treatment levels (program control, learner control, and learner control with advisement) and three within-treatment levels (pre-task, on-task and post-task state anxiety levels) was used to examine state anxiety measures. For achievement, a pretest-posttest experimental design was used. Subjects randomly assigned to the three treatment conditions were administered the same pretest and posttest measures along with periodic administrations (SA1, SA2, and SA3) of the S-Anxiety subscale of the STPI before, during and after instructional treatments. A multiple linear regression model was used to test the relationship between posttest achievement and pretest performance, on-task state anxiety level (SA2), treatment, gender, and the two-way interactions of treatment by pretest and treatment by on-task anxiety.

Results

Anxiety-Treatment Interaction

Means and standard deviations for the group-treatment factor and within-subject factor are presented in Table 1. The split-plot analysis of variance of state anxiety change across measurement occasion for treatment conditions, shown in Table 2 demonstrates that there was no significant difference for the main effect of treatment, $F(2,103) = 0.40$, $p > .05$. In addition, there was no significant interaction between treatments (program control, learner control and learner control with advisement groups) and occasion (pretask, on-task and post-task) for state anxiety levels. The computed F ratio for the two-way interaction: Occasion by Treatment was $F(4,206) = 1.46$, $p > .05$.

There was a significant main effect for the within-subject variable of measurement occasion, $F(2,206) = 9.17$, $p < .05$. Specifically, this significant within-subject difference in state anxiety levels across measurement occasions reflects a general reduction of subjects' perceived level of state anxiety as the CAI lesson proceeded.

Insert Table 1 About Here

Insert Table 2 About Here

Independent Variables and Posttest Performance

A multiple linear regression model was used to test the relationship between posttest achievement and pretest performance, on-task state anxiety level (SA2), treatment, gender, and the two-way interactions of treatment by pretest and treatment by on-task anxiety. Results of the analysis are presented in Table 3. The regression analysis yielded a significant relationship, $F(1,96) = 6.93$, $p < .05$, between pretest scores and posttest scores following the CAI treatments. Pretest scores are a reliable predictor of posttest performance by study subjects.

The other independent variables of on-task anxiety, instructional treatment and gender revealed no significant relationship between the variable and posttest achievement. The computed F ratios for each of these independent variables are presented in Table 3.

It was also found that there was no significant relationship between the interaction of CAI instructional treatments and pretest scores on posttest achievement. In addition, there was no significant relationship between the interaction of CAI instructional treatments and on-task state anxiety levels on posttest achievement (Table 3).

Insert Table 3 About Here

Time

Means and standard deviations for the time required for each treatment to complete the CAI lessons are presented in Table 4. A one-way ANOVA yielded a significant main effect, $F(2,103) = 4.48$, $p < .05$, for treatment groups in the time required to complete the CAI learning task. Tukey's post hoc analysis indicated that a significant difference existed between the learner control with advisement treatment condition and the learner control condition with regard to time required to complete the CAI lesson. The learner control group used significantly less time to complete the CAI lesson than the learner control with advisement group.

Insert Table 4 About Here

Discussion

The results of this study indicate that providing learners with control or control with advisement over the instructional variables of lesson sequence, number of practice questions to answer, review of content related to questions answered incorrectly, and the decision when to terminate the CAI lesson had no significant impact on learner anxiety levels when compared to a program controlled CAI design. Three earlier CAI studies have previously reported that providing learners with control of instructional variables did result in decreases in state anxiety levels (Collier et al., 1973; Gallagher et al., 1971; Hansen, 1972). However, the reported decreases in state anxiety were not significantly lower than similar treatments where learner control was not provided. The present study similarly indicated that utilizing various locus of instructional control options does not produce differing levels of state anxiety that may in return impact learning. Therefore, locus of instructional control strategies may not be as important a factor in determining learner state anxiety levels as previously believed.

There was no interaction between CAI locus of instructional control strategy and occasion for measurement of state anxiety levels. The state anxiety measurements of study subjects indicated a similar gradual reduction of state anxiety for all treatment groups across the CAI lesson. A number of researchers have theorized and tested the hypothesis that providing learners with control over certain instructional variables would enhance learning and reduce anxiety levels (Collier et al., 1973; Gallagher et al., 1971; Hansen, 1972; Judd et al., 1975). This study supported the results of Hansen (1972) and Judd et al. (1975) in finding that learner control (or the learner control with advisement) had no different effect on state anxiety levels in instructional situations than caused by an imposed control (program control) condition. The program control condition did not create higher anxiety levels that may have indicated that learners felt constrained or overly directed by the imposed control.

Several possible explanations may have accounted for the significant decreases in state anxiety levels across measurement occasions for the treatment groups. First, all learners, when exposed to a new, unfamiliar and potentially threatening learning experience can be expected to initially exhibit high state anxiety levels in response to the situation. In this study learners may have become familiar and more comfortable with the learning situation and overcame their initial state anxiety reaction.

A second reason may have been that many individuals perceive any sort of technology as potentially threatening. The emergence of computers in education has led to concerns, both cognitively and emotionally based, about the response of learners to the new technology (Cambre & Cook, 1985). Computer anxiety has been conceptually defined as "the mixture of fear, apprehension, and hope that people feel when planning to interact or when actually interacting with a computer (Rohner & Simonson, 1981, p. 551)." Since a majority of subjects in this study (61%) indicated that they had little or no computer experience, it could be expected that they would react initially with high levels of state anxiety. In general, the CAI treatments resulted in no negative effects on learner anxiety levels or attitude. Upon completing the CAI lesson, students were asked to rate their attitude (negative, neutral, or positive) toward the CAI lesson. Results of the attitude survey indicated that 86% of all study subjects responded positively to the CAI learning experience, 14% indicated that their feelings were neutral, while no one responded negatively toward the CAI lesson.

The results of the multiple linear regression analysis indicated that there was a significant relationship between study subject's pretest scores and their posttest scores following the CAI treatment. For the CAI lesson, a learners' pretest score appears to be a reliable predictor of their performance on the posttest measurement. By establishing this relationship between prior knowledge and posttest achievement, CAI designers can use pretest scores to identify those learners requiring extra instructional support.

The regression analysis indicated that there was no significant relationship between subjects' on-task (SA2) state anxiety levels and their posttest performance. In the present study, subjects who were high in state anxiety and subjects who were low in state anxiety did not differ with regard to posttest achievement resulting from the CAI treatments. Tobias (1979) has suggested that high anxiety can be debilitating in instructional situations. However, this study indicated that state anxiety levels did not affect learning performance.

The CAI lesson designs may have been a more important factor in assuring that anxiety or lack of anxiety had a minimal impact on learning. In the design of the CAI courseware, recommendations were followed from Tobias' (1979) model which provides instructional researchers with guidelines for designing instruction to minimize anxiety in instructional situations. The provision for immediate practice of instructional content, remediation, well organized content, and providing students with the ability to reexamine instructional input (available only for learner control and learner control with advisement groups) may have minimized any influence that anxiety may have had during the learning task.

The data analysis indicated that there was no significant relationship between CAI treatment group and posttest achievement. The mean posttest scores were 18.94 for the program control, 17.97 for the learner control and 18.66 for the learner control with advisement treatment group. While there was no significant posttest achievement difference between the three locus of instructional control strategy treatments, subjects of the three treatment groups did answer an average of 77% of the posttest questions correctly.

A number of researchers who have reviewed the learner controlled research have indicated that learner control strategies do not result in any drop in learning, yet there was no clear indication that the overall use of learner control caused any increase in student achievement, efficiency, or attitude toward instruction (Merrill, 1980; Milheim & Azbell, 1988; Steinberg, 1977). This research study is supportive of these conclusions that learner control possesses no advantage over imposed control instruction.

Clark (1984) and Carrier (1984) suggested that learners may not be sophisticated enough to effectively direct their own learning strategies when provided control over instructional variables. Hannafin (1984) suggested that providing learner advisement during the instructional task would assist students in utilizing the learner control features more effectively. The results of this research study however, do not support the recommendations made by Hannafin (1984). In the present study, the advisement provided to learners in the learner control with advisement group did not result in any greater achievement on the posttest when compared to posttest achievement of the learner control group which did not receive advisement. The addition of the advisement to the learner control strategy actually increased the instructional time utilized by the learner control with advisement group, yet there was no associated learning gain.

Several important implications for educational researchers and CAI courseware designers are provided by this study. The three locus of instructional control strategies implemented in this research study did not have any differential effect on learning. Therefore, the issue of whether to incorporate differing locus of instructional control options in CAI courseware designs may not be as critical as previously believed. This study does suggest that well planned systematically designed courseware can be effective when certain guidelines are followed. In this study, all three CAI treatments consisted of well organized instruction, individual student pacing of the lesson, embedded practice questions, immediate feedback, and opportunities for remediation. Computer-assisted instruction designers should be aware that using a systematic approach utilizing a model such as the one presented by Roblyer (1988) may be more important than locus of instructional control considerations in the design of CAI.

In this study it was determined that there was no significant relationship between a subjects gender and their posttest performance. Although the computer experience level for males and females were quite different, it did not affect learning performance. The computer experience level of males was considerably higher than for females. Before the computer lesson, students were asked to indicate their self-perceived computer experience level (a. little or no computer experience; b. moderate computer experience; or c. extensive computer experience). About 46% of males compared to 70% of females indicated that they possessed little or no computer experience. About 51% of males reported having moderate computer experience while only 28% of females reported the same experience level. Only 3% of males indicated they had extensive computer experience compared to less than 2% for females. With this large disparity between computer experience levels of males and females, one might expect some effect on learning. However, no effect was evident. Any gender or differences in computer experience levels of the subjects was apparently counteracted by the CAI lesson design.

An interaction between pretest scores and CAI treatments on posttest performance was not found. This is in contrast to a study by Gay (1986) who investigated learner control and a student's prior conceptual understanding in a computer-assisted interactive video lesson. Gay found that low aptitude subjects in the learner control group had significantly lower posttest scores than the low aptitude subjects in a program control group and lower posttest scores than high aptitude subjects in either group. The lack of an interaction in the present study suggests that learners of all entering ability levels were able to effectively use the control options provided.

There was no significant relationship between the interaction of treatment and on-task state anxiety levels on posttest achievement. No treatment condition served to provide learners, with varying levels of anxiety, with a more efficacious lesson design strategy for affecting posttest achievement. State anxiety levels may therefore not be an important consideration when designing CAI courseware that will utilize various locus of instructional control methodologies.

The time required to complete the CAI lesson was significantly less for the learner control group than required by the learner control with advisement group. This was the only significant difference among the three groups with regard to time. While the learner control with advisement group utilized 4.62 minutes more than the learner control group, there was no associated posttest achievement gain. This finding is similar to the study by Goetzfried and Hannafin (1985) who also found no achievement gains resulting from increased learning time required by learner control with advisement group. In this study students in the learner control condition seemed to be a good judge of when to answer embedded questions and when to review the content related to the questions in the case of a wrong answer. Students in the learner control condition answered an average of 20.9 questions while the learner control with advisement group answered an average of 30.9 questions (students in the program control condition answered all 29 embedded questions). This is a very critical issue when dealing with valuable instructional time and in designing CAI courseware.

The CAI designer must be aware that the various locus of instructional control strategies require different amounts of instructional time for students to complete the lesson. In this study, the learner control with advisement treatment required significantly more time to complete the lesson than did the learner control condition yet showed no significant achievement gain. Students in the learner control with advisement group not only spent more time on task but also attempted more embedded questions as well. The learner control with advisement lesson apparently created greater diligence to the learning task. Additional research would be useful in determining the most effective use of advisement and in what situations advisement should be provided. Finally, it would also be valuable to examine if the extra time on task required by the learner control with advisement group impacted long-term retention of the CAI content when compared to the other lesson design strategies. The extra time used by this group may ultimately impact long-term retention and be worth the additional short-term effort.

References

Cambre, M.A., & Cook, D.L. (1985). Computer anxiety: Definition, measurement, and correlates. *Journal of Educational Computing Research, 1*(1), 37-54.

Carrier, C.A. (1984). Do learners make good choices? *Instructional Innovator, 29*(2), 15-17.

Clark, R.E. (1984). Research on student thought processes during computer based instruction. *Journal of Instructional Development, 7*(3), 2-5.

Collier, R., Poyner, L., O'Neil, H.F., Jr., & Judd, W.A. (1973, February). *Effects of learner control on performance and state anxiety in a computerized learning task.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Fennema, E.H., & Sherman, J.A. (1977). Sex-related differences in mathematic achievement, spatial visualization, and affective factors. *American Educational Research Journal, 14,* 51-71.

Gallagher, P., O'Neil, H.F., Jr., & Dick, W. (1971, February). *Effects of trait state anxiety in a computer-managed instruction course.* Paper presented at a meeting of the American Educational Research Association, New York.

Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. *Journal of Educational Psychology, 78*(3), 225-227.

Goetzfried, L., & Hannafin, M.J. (1985). The effect of the locus of CAI control strategies on the learning of mathematics rules. *American Educational Research Association, 22,* 273-278.

Hannafin, M.J. (1984). Factors affecting the locus of instructional control in the design of computer-assisted instruction. *Journal of Instructional Development, 7*(3), 6-10.

Hansen, J.B. (1972). *An investigation of cognitive abilities, state anxiety, and performance in a CAI task under conditions of no feedback, feedback and learner control.* Austin: University of Texas Computer Assisted Instruction Laboratory. (ERIC Document Reproduction Service No. ED 072 634).

Judd, W.A., Daubek, K., & O'Neil, H.F. (1975, March). *Individual differences in learner controlled computer-assisted instruction.* Paper presented at the annual convention of the American Educational Research Association, Washington, D.C.

Leherissey, B.L., O'Neil, H. F., Jr., & Hansen, D.N. (1971). Effects of memory support upon anxiety and performance in computer-assisted learning. *Journal of Educational Psychology, 62,* 413-420.

Louie, S., Luick, A.H., & Louie, J.L. (1985). Locus of control among computer using school children: A report of a pilot study. *Journal of Educational Technology Systems, 14*(2), 101-118.

Merrill, M.D. (1980). Learner control in computer based learning. *Computers and*

Education, 4, 77-95.

Milheim, W.D., & Azbell, J.W. (1988). *How past research on learner control can aid in the design of interactive video materials.* In Proceedings of the Annual Meeting of the AECT. (ERIC Document Reproduction Service No. ED 295 652).

O'Neil, H.F., Jr., (1972). *Anxiety reduction and computer-assisted learning.* Paper presented at a meeting of the American Psychological Association, Honolulu, Hawaii.

O'Neil, H.F., Jr., Spielberger, C.D., & Hansen, D.N. (1969). The effects of state anxiety and task difficulty on computer-assisted learning. *Journal of Educational Psychology, 60,* 343-350.

Roblyer, M.D. (1988). Fundamental problems and principles of designing effective courseware. In D.H Jonassen (Ed.), *Instructional designs for microcomputer courseware.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Rohner, D.J. & Simonson, F.R. (1981, April). *Development of an index of computer anxiety.* Paper presented at the annual convention of the Association of Educational Communication and Technology, Philadelphia, (ERIC Document Reproduction Service No. 207 487).

Ross, S.M. (1984). Matching the lesson to the student: alternative adaptive designs for individualized learning systems. *Journal of Computer-Based Instruction, 11*(2), 42-48.

Ross, S.M., & Rakow, E.A. (1980). Adaptive design strategies for the teacher-managed course. *Journal of Instructional Psychology, 7,* 13-19.

Seymour, S.L., Sullivan, H.J., Story, N.O., & Mosely, M.L. (1987). Microcomputers and continuing motivation. *Educational Communication & Technology Journal, 35*(1), 18-23.

Sieber, J.E., O'Neil, H.F. Jr., & Tobias, S. (1977). *Anxiety, learning, and instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Snow, R.E. (1980). Aptitude, learner control, and adaptive instruction. *Educational Psychologist, 15,* 151-158.

Spielberger, C.D. (1966). *Anxiety and behavior.* New York: Academic Press.

Spielberger, C.D. (1979). *Preliminary manual for the state-trait personality* *inventory*. (STPI). University of South Florida.

Steinberg, E.R. (1977). Review of student control in computer-assisted instruction. *Journal of Computer-Based Instruction, 3*(3), 84-90.

Tennyson, R.D. (1980). Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. *Journal of Educational Psychology, 72*(4), 525-532

Tennyson, R.D. (1981). Use of adaptive information for advisement in learning concepts and rules using computer-based instruction. *American Educational Research Journal, 18*(4), 425-438.

Tennyson, R.D., & Buttrey, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. *Educational Communication & Technology Journal, 28*, 169-176.

Tobias, S. (1979). Anxiety research in educational psychology. *Journal of Educational Psychology, 71*(5), 573-582.

TABLE 1.

Means and Standard Deviations for State Anxiety Measurement Occasions by Treatment Condition

Control Strategy		SA1	SA2	SA3
Program Control	Mean	18.05	16.38	15.84
(N=37)	SD	(5.00)	(4.90)	(5.32)
Learner Control	Mean	16.67	15.65	15.35
(N=37)	SD	(5.20)	(5.73)	(5.27)
Learner Control	Mean	16.12	16.41	15.19
Advisement	SD	(4.58)	(4.98)	(5.10)
(N=32)				

Notes: SA1 = State Anxiety Measurement 1 (pre-task); SA2 = State Anxiety Measurement 2 (on-task); SA3 = State Anxiety Measurement 3 (post-task).

TABLE 2.

Source Table of ANOVA of State Anxiety Change Across Measurement Occasions for Treatment Conditions

Source	SS	df	MS	F
Between Treatments				
Treatments	53.23	2	26.62	0.40
Error	6847.29	103	66.48	
Within Treatments				
Occasion	117.80	2	58.90	9.17*
Occasion X Treatment	37.64	4	9.41	1.46
Error (occasion)	1323.40	206	6.42	

*p< .05 with Geisser-Greenhouse correction

TABLE 3.

Linear Regression to Test the Relationship Between Independent Variables and Posttest Performance

Source	Type III SS	df	MS	F
Pretest	76.55	1		6.93*
SA2	15.17	1		1.37
Treatment	5.22	2		0.24
Gender	1.43	1		0.13
Pretest X Treatment	1.99	2		0.09
SA2 X Treatment	37.22	2		1.69
Error	1060.31	96	11.04	

*$p < .01$ Note: The overall test of significance was $F = 1.85$ with $df = 9,96$.

TABLE 4.

Means and Standard Deviations for Time to Lesson Completion by Treatment Condition

Control Strategy	Time (in minutes)	
Program Control (N=37)	Mean SD	26.88 (6.48)
Learner Control (N=37)	Mean SD	23.75 (6.97)
Learner Control Advisement (N=32)	Mean SD	28.37 (6.26)

Title:

Using Photo-Interviewing as Tool for Research and Evaluation

Authors:

John V. Dempsey
Susan A. Tucker

Introduction

The tradition in the area of instructional technology has been consistent in its approach to research and evaluation. The focus has been on documenting the success and failure of objectives. To a large extent, qualitative dimensions have been avoided. Unfortunately, in numerous research and evaluative efforts (especially those that are internally generated), objectives are not always clearly defined. Even those that are clearly defined in the proposal often suffer Murphy's law. Particularly when a program is in its first year of operation, reductionistic and objectives-based data collection strategies generate insufficient data for ongoing program improvement. Increasingly, instructional designers and educational psychologists are turning to innovative methods to collect and analyze data for interpretive research studies as well as naturalistic or improvement-oriented evaluations.

For these reasons, we propose that one other data collection methodology be added to the "standard fare". Photo-interviewing (Tucker & Dempsey, 1990) was selected as the strategy because it offers much promise in collecting holistic and consistent samples or records of critical phenomena. The viability of this method has been well documented in anthropology and sociology (e.g., Becker, 1978; Byers, 1964; Collier & Collier, 1986; Garfinkel, 1967) and has been further illuminated by Susan Sontag's (1977) cautions. Even so, photography as a method of inquiry has not been applied systematically in educational research or program evaluation (Fang, 1985; English, 1988). We need to know more about the rhetoric of words and pictures in educational inquiry. To unlock the meaning of this information, one must examine more than empirical facts. We need to grapple with the problem of gathering subjective and intersubjective perceptions. Central to our investigation is the study of what Erickson & Mohatt (1982) call "immediate environments of learning", namely an analysis of how instructional contexts are socially and psychologically constituted.

According to Collier & Collier (1986), "the production of photographs for purposes of research requires ...an overview of the culture or setting which allows all details to be seen within a context." This whole view of a context can be achieved in a variety of ways. This paper will focus on some of the techniques and challenges of using photographic evidence and photo-interviewing in educational evaluation and research contexts. For the purpose of this paper, we will differentiate the two terms on the basis of five dimensions: purposes, role of the client, methods, controls, and ends. *Research* aims for generalization, gathering descriptions, seeking conclusions and trying to explain causation. Directed by the researcher, research's methods often emphasizes the empirical, seeks internal and external validity as its controls, and the ultimate results may or may not have an immediate payoff. Contrast these characteristics with evaluation. *Evaluation* aims for satisfying specific situations and synthesizes judgments from a myriad of data sources in order to facilitate decisions, rather than explaining causes. Directed by the client, evaluation's methods are multi-disciplinary, and uses utilitarian versus statistical controls (e.g., accuracy, utility, feasibility, and credibility). The ultimate results are expected to be immediate and practical.

Why Photo-Interviewing?

We believe that photo-interviewing yields richer data than that usually obtained from verbal interviewing procedures alone. Informants tend to examine images and react to cues present in those images more carefully than would have been expected using written or spoken cues alone. Photographs trigger recall and focus the interviewing process, enabling an in-depth look at intended as well as unintended aspects of a program. Photographic images solicit both differences and similarities in individual perception which can be validated and analyzed across groups of perceivers.

Photo-interviews can generate data particularly helpful for pilot-stage projects where innovation and development rather than status quo and maintenance are organizational priorities. Moreover, we posit that photo-interviewing can serve as an instructional as well as evaluative strategy given the projective nature of visual images. Photographs prompt reflection upon the program being evaluated which goes far beyond an interviewing situation lacking visual cues.

Photographic evidence of behavioral patterns can be "detected" in many ways. In a classroom environment, for example, visual data can include student body language, seating patterns, and off-task behavior. Photographs appear to act both as stimuli and verifiers of perception. By providing a common stimulus for informants, photographic images both solicit differences and similarities in individual perception which can be analyzed across groups of perceivers.

Moreover, a particularly potent argument for using photography (including film or video) is the potential for multiple applications in evaluation and research situations. Photographs capture direct observations of an environment and can serve as stimuli for interviews which in turn generate more data. In this sense, photography and its related media are particularly versatile in those environments in which methodological triangulation is applicable. This argument also raises the valid counterargument that using photographs as data and stimuli, for example, may lead to inconsistent or contradictory conclusions. This really is the argument against using any form of triangulation. Mathison (1988) asserts that it is unrealistic to expect that the use of multiple data sources, investigators, or methods can lead to a singular proposition about the phenomenon being studied. While his rationale may work from a behaviorist tradition, cognitivist models of reality seek holistic data collection, risking the price of inconsistent or even contradictory data in their quest for the wealth of information derived from multi-modality data sources. Admitting the inconsistencies and contradictions which must arise in using photo-related data sources is not meant to justify an absence of some theoretical perspective. Rather, we contend that it is only through the use of robust cognitive theories that a researcher or evaluator has hopes of constructing plausible explanations of diverse phenomena.

Episodes: Applications in education and training

Our ideas concerning applications or "episodes" of photo-interviewing are based on two of our attempts to incorporate photo-interviewing with inquiries involving separate program innovations. Episode 1 involved an evaluation of an intensive 40-hour hypermedia mini-course for higher education faculty and graduate students. The setting of the episode was a computer lab. Episode 2, which is currently in the process of analysis, involved research with high school juniors and seniors and took place in a chemistry classroom. The study considered ways that matrices can function as systematic organizing structures for concept and simple rule learning. In the discussion of the photo-interviewing process which follows, we will use examples from both of these inquiries.

Photo-Interviewing: The Process

The function of all analysis is to find patterns. To identify patterns, photo-interviewing uses photographic tracking and selective responses to "peaks of information", when and where they happen. This is a flexible approach which can collect unpredictable as well as scheduled behaviors in research and evaluation contexts. As Hall (1974) advises, we approach the photographic content with an open strategy in order to respond to their holistic content. At present, the strategy we propose consists of a nine-step process.

1. Identify major questions and research or evaluation perspective.

The purpose of this stage is to clarify research or evaluation questions and approaches.

Example from Episode 1: An evaluation paradigm was negotiated by evaluators with stakeholders concerned with the hypermedia workshop. Consistent with initial discussions, this negotiation process delineated major questions, sources of evidence, and quality standards for the evaluation. Major questions were divided into those affecting the context, process, and product aspects of the evaluation.

Context questions dealt with four issues. Why was the class planned? How was the class planned in terms of: goals, objectives, and priorities, staffing, methodologies, prerequisites, and assessment? What characterized the setting in terms of unique or common factors as well as student and faculty backgrounds? Finally, what resources are available (e.g., hardware, software, text, readings, lecture notes, overheads, films, and guest lecturers)?

Process questions first considered how ongoing events evolved. What were student, faculty, and trainer perceptions over time? What was the planned versus real content? What were the barriers to success? Also an important process consideration centered on effectiveness. How effective were various events which took place during the mini-course, (i.e., resources, processes,handouts, readings, cooperative learning events, cognitive versus behavioralistic teaching methods)? Which of Gagne's events of instruction were employed during particular periods? How receptive were the trainees at different junctures? What hindered and helped? What were the "light bulb events"?

Product questions began with how well the predetermined objectives were met. a related question considered which consequences (intended and unintended) were actually attained. We were also concerned student perceptions regarding the degree of match with early expectations. Other questions negotiated dealt with the transfer value of content and processes. Product-related questions in this particular evaluation were intended to focus especially on prerequisite skills, content, processes and sequencing.

Sources of evidence included photographs of the mini-course as well as participant comments from photo-interviewing sessions. Many other sources of evidence were collected as part of the negotiated paradigm. These include the course syllabus, context inventories, daily benchmarks, class size and attendee mortality, faculty vitae, student records, trainee demographics, evaluator notes, summative questionnaires, and HyperCard products produced by students.

Quality standards attempted to match context, process, and product evaluation questions. These were numerous, but included depth and breadth measurement; generating information that both documents and leads to program improvement; ensuring the transferability of content and processes; matching expectations with actual outcomes; communicating well-defined pre-requisites and accommodating diverse disciplines; the attaining both intended and unintended competencies; using the events of instruction effectively; efficient sequencing of the course; supplying adequate resources; integrating motivation throughout the course.

Comment: Visual anthropologists and visual sociologists recommend using a less structured approach toward inquiry when first examining natural phenonema. Program evaluation is much more focused in terms of major questions from the very beginning and this is reflected by our paradigm. Since photography is negotiated along with other sources of evidence and the aim is to lead toward program improvement, we designed our inquiry more deliberately than anthropologists or sociologists.

2. Prepare the data-gathering approach.

The purpose of this stage is to establish content parameters for the photographs (i.e., to isolate topics related to critical incidents). An additional task was to decide on the frequency and extent of coverage.

Example from Episode 2: After a good deal of contextual analysis, the researchers decided to concentrate on five topics related to the descriptive research. These topics were describing the strategies used in learning, group work and partnering, feedback among participants, teacher contact, and reward-related occurrences. The photographer was free to take as many photographs as possible as long as he didn't feel he was being obtrusive. Based on prior experiences, such as Episode one, the researchers decided that photographs would be taken at the discretion of the photographer rather than at set (predetermined) intervals.

Comment: From one perspective, it is important that photographers functioning as data collectors are extremely open and receptive to the "unanticipated" in the environment in which they work. On the other hand it is also important that photographers have a perspective and purposely capture the "intended" in their inquiry. An excellent photojournalist, for example, may be assigned to cover a fire. The photographer knows she must "capture" critical events or incidents related to topics in newspaper or magazine readers' "fire" schemata. The topics are defined, yet the photographer is free to interpret in innumerable ways.

The following quote is attributed to John Szarkowski, the influential curator of the New York's Museum of Modern Art.

> Photography is a system of visual editing. At bottom, it is a matter of surrounding with a frame a portion of one's own cone of vision, while standing in the right place at the right time. Like chess, or writing, it is a matter of choosing from among given possibilities, but in the case of photography the number of possibilities is not finite but infinite. (Sontag, 1977, p. 192).

By establishing topics which later will be connected with critical incidents used in photo-interviewing, the photographer/data collector becomes both clear about the parameters of the episode and responsive to unanticipated phenomena.

3.Photograph episode.

The purpose of this stage is to gather visual data describing the environment.

Example from Episode 1: The mini-course was divided into two Friday-Saturday-Sunday segments. This was one of our first experiences with formative evaluation using photo-interviewing and we were following the pioneering work of Collier (1967). Thus, we opted during the first weekend to shoot at predetermined time intervals (every 20 minutes-- one close-up, one wide angle). After reviewing the effect of this strategy, we decided on a more participant-observer approach for the second three-day session which responded to critical incidents rather than time. Also, during the first three-day session, we used an electronic flash for illumination. During the second session, available light was used as a light source. When interviewed later, not one trainee (n=27) observed any differences between the first three-day session and the second. Further, most participants forgot about being photographed after the first 1/2 hour of the course.

Comment: In photographing program participants as part of an inquiry, one is reminded of Ziller's (1990) experiments with Self Theory in which subjects were given photographic equipment and instructed to photograph various abstract concepts. There is just no separating the photographer's self from the photographer's observations. The question is should a photographer intent on gathering the most useful data attempt to record "the decisive moment" as the work of Cartier-Bresson (1952) exemplifies or, alternately, map the environment as a type of social cartography. In responding to this question, first consider the nature of detection. Sometimes the camera, as a powerful instrument of description, should be called on to interpret diverse perspectives. On other occasions, the camera is needed only to record. Given our research in natural settings, we propose, photographers should consciously operate dualistically--both as an interpreter and as a recorder. It may even be argued that artificially imposing constraints such as shooting only at specified time intervals or using "standard" camera angles or lenses actually contributes to a most serious threat to validity: poor visual description.

4. Organize and label photographs.

The purpose of this stage is to organize photographic data in a fashion that will be most useful for photo-interviewing.

Example from Episode 2: Photographs (shot in black and white) were processed and 3 1/2 X 5 inch proofs were made of each negative. The proof prints were numbered sequentially by roll (e.g., B-31 would be the photo ID number for the 31st exposure on roll B). The classroom clock was included in the photographs every so often during the study to provide a temporal reference to accompany the chronological succession of images. The time for various shots could also be approximated by refering to the stationary video cameras of the classroom that were always running during the study.

After the photos were labeled, the researchers viewed the data as a whole. Photographs were screened individually and then, collectively. Data collectors and evaluators interviewed each other about perceptions invoked by the photographs in an effort to balance interpretative with recording shots. Questions and verbal comments of the researchers were noted and referenced to photo ID numbers.

After participants were scheduled for small group interviews (typically triads), photographs were clustered to include specific shots of individuals within a particular collaborative learning groups being studied as well as generic shots of the whole class. Finally, the photographs were organized to correspond to topics complementing the five research questions under study.

Comment: Whatever the strategy employed, photographs should be assembled in preliminary structured units which document a visual narrative. For example, photo events can be arranged as chronological experiences, as spatial maps of an environment, or as models.

In effect, the analysis begins here. This stage roughly conforms to the beginning of Collier & Collier's basic model for analysis (1986, p. 178). At the same time, the stimuli is arranged in a way that is intended to provide for efficient interviewing. It may be necessary, for example, to cull individual photographs which are clearly redundant. In our first attempts at photo-interviewing, we overwhelmed some interviewees by placing too many photographs on the table for their consideration. In trying to mentally sort the photographs into logical catagories for discussion, they became confused and often "got off-track" in their conversation. After that experience, it was clear to us that it is the researchers responsibility to limit the stimuli in a way that both helps subjects' recollections and allows some choice among the photographs selected.

5. Prepare protocol.

The purpose of this stage is to prepare a protocol (i.e., of verbal and visual questions) for the photo-interviewing session.

Example from Episode 2: These oral interview questions were organized around the evaluation paradigm's five major questions which concentrated on learning strategies, group work and partnering, feedback, teacher contact, and rewards. In this study, we generated a verbal interview schedule which explored each research question and then reorganized the question sequence using a loose adaptation of the critical incident technique.

The protocol or script was very specific in its wording so that the same questions were asked of all cooperative learning groups. After a brief statement in which subjects were informed that the researchers were asking the interview questions to see how well some methods of learning work, subjects were asked for general reactions to the five topics. The reader is referred to Appendix A for a delineation of the interview schedule.

The remainder of the protocol, with the exception of directions, were phrased in terms of critical incidents. A section of the interview script, concerned with learning strategies, is shown in Appendix B.

Comment: Flannagan (1954) defines critical incidents as specific reports of observed behavior from qualified sources. The critical incident technique is remarkable in that it provides open-ended responses which tend to avoid generalizations or opinions. This technique, which began as a military training invention, is used in job analysis for those tasks with a flexible or undefinable number of "right" ways to behave (Zemke & Kramlinger, 1985). In our judgment, this technique encourages participant sharing and the emergence of perceiver diversity without the typical oppression of the conforming group.

6. Pilot interview.

The purpose of this stage is to go through a "dry run" of the photo-interview.

Example from Episode 2: In this study, as a result of last minute schedule changes at the school site, we were able to pilot the photo-interviewing session only with members of the research team and the class instructor. As such, we were primarily interested in what could go wrong with the photo-interviewing session. Were all photos numbered correctly? Was the video equiptment which would record the session properly checked-out? Could any questions in the interviewing script be made clearer? Was the interview text clearly associated with the research questions? Could the photo-interview be completed in the time we had available? Several items in the interview were revised and some photographs culled as a result of this process. This process also resulted in what we referred to as a "bring along checklist" which detailed the procedures, timelines, materials, and equiptment necessary for conducting multiple photo-interviews.

Comment: Piloting the photo-interview with individuals or small groups of the target population can minimize implementation problems. While we did not encounter major problems, Our efficiency could well have been improved by conducting one-to-one evaluations (Dick & Carey, 1990) with a sample of high school science students. Piloting with target members ensures that the interview questions, photographs, and procedures are age and ability appropriate. Converse and Presser (1986) recommend ten purposes for "pretesting" the instrument. The first four evaluate specific questions for variation, meaning, task difficulty, and respondent interest and attention. The last six bear on the instrument as a whole. These weigh the "flow" and naturalness of the sections; the order of the questions; skip patterns; timing; respondent interest and overall attention; and respondent well-being.

7. Conduct and videotape interview.

The purpose of this stage is to elicit and collect a flow of information regarding the research or evaluation questions.

Example from Episode 2: The photo-interviews closely followed the protocol or script discussed in step 5. After the interviewer reviewed the purpose of the study and made assurances of confidentiality, four groups of three subjects were interviewed. (There were several absences due to illness.) As this study considered two separate concept training strategies, two photo-interview sessions were held with each group. Interviews were scheduled a few days after each training session. Photographs discussed during this procedure were used to clarify and extend the data derived from the oral interviews.

In each interview, there was one interviewer. A videocamera with a remote microphone recorded the interviews for later analysis. A second photo-interviewer was also in the background, taking notes and occasionally asking probing questions.

Comment: During episode 2 we used critical incidents along with photographs. This is somewhat different than other researchers using photo-interviewing. Probably this approach is not useful for all situations. Even so, there are three advantages in using questions aimed toward critical incidents. First, this approach encourages the perspective of "open- inquiry" which is particularly valuable in innovative development projects. The subject is encouraged to respond or project as they see fit. Second, questions phrased in this way get at facts as perceived by the subject-observer. They encourage description, not judgements. The description, however, is open to subject projections. Using photographs and anecdotes, the subject has the evidence to describe to the interviewer, what in his or her opinion "really happened". Finally, critical incident responses are focused on specific observable occurrences related to the research or evaluation. The critical incident technique works with the photographs to assure that subjects' responses spotlight scenarios related to reseacher or evaluation questions without cuing subjects to a particular viewpoint.

8. Analyze photographs, field notes, interview notes and videotape.

The purpose of this stage is to digest visual, auditory, and written records in ways that provide insight and information.

Example from Episode 1: Visual data (still and video) are supersaturated with information. This enormous quantity of information presents great opportunity and even greater difficulty in approaching analysis. In episode 1, we were faced with making sense of a particular class that was, after all, one small part of a continuing three year project involving very diverse sources of evidence. .

Collier and Collier's (1986) basic model for analysis was adapted for our particular needs. We began by organizing the data somewhat sequentially. We kept a fairly a narrative log of the visual chronology, spatial relationships of participants, and instructional content and processes during the course. We also had detailed records of the photo-interviewing process. These we gathered along with context and summative questionnaires, stills, and video to provide a context for our inventory. As much as possible, we organized our data into informal categories related to our research questions. Our approach at this point was what Collier and Collier refer to as "open viewing".

> Open viewing is an unstructured immersion in the visual record, a repeated viewing of all the material that allows you to respond to the images as they are and not simply as you expect them to be. (1986, p. 181).

To approach structured analysis, we returned to the paradigm negotiated at the beginning of the evaluative effort and especially to the major questions identified in the first step of this process. Context, process, and product questions provided a basis for organizing our perceptions. Although we did not reduce the questions to the level of microanalysis, we focused the data analysis by defining the evaluation questions and issues related to these questions.

Throughout both the open viewing stage and the focused analysis stage, notes were taken by the evaluators and much discussion ensued regarding findings. Collier and Collier (1986) have commented on the value of these exchanges between researchers.

"These discussions clarified details, raised important questions, and defined conclusions,and the interplay of ideas sharpened our examination of the evidence and the precision of our analysis. On a conceptual level these joint viewings were the most productive stages of the research. " (p. 177).

Comment: Finding an effective approach to analysis was (and is) the most difficult part of this process for us. Our approach did not include microanalysis. Our review of the literature uncovered three behavior patterns which constitute evidence for microanalysis: proxemics (measurements of space); kinesics (messages of body behavior); and choreometrics (patterns of behavior through time). Like many evaluators and researchers, we were struggling to control our basic techniques for descriptive analysis without concentrating on a purely "count, measure, and quantify" approach. We felt that quantifying visual information, while being easier to proceduralize, would tend to greatly lengthen the analytical process with no guarantee that the results would produce meaningful conclusions. On the other hand, we certainly do not disdain the use of quantitative microanalysis. We only agree with Collier and Collier (1986), that microanalysis is a reductionist, clinical technique useful in studying small components of human behavior. Particularly in the case of naturalistic evaluation, microanalysis is not compatible with the paradigm.

9. Summarize findings.

The purpose of this step is to make sense of the data analysis and offer conclusions to interested parties and stakeholders.

Example from Episode 1: Evaluation findings were used for internal and external purposes. Both uses entailed a two step editing procedure.

First, the findings were used as part of the internal evaluation of the course, the aim being to generate *program improvement* strategies. A written summary was submitted to project staff. This report was accompanied by annotated "generic yet contextually rich" photographs included in the body of the text to highlight major findings. Sensitive findings as well as "particular" photographs which could be traced back to a particular individual were shared verbally in a confidential setting but not put to paper. This first report attempted to validate findings and avoided the inclusion of judgments. Afterwards, we edited the initial findings, incorporated revisions generated by reactions to the first "draft" and added an implications and judgment section.

The findings were also used as part of the annual *documentation* to the funding agency which supported the development of the hypermedia course. Cleansed by the internal evaluation reporting process described in the paragraph above, a "final" annual summary was submitted to the funder. Using an executive summary format, it also incorporated generic photographs (though fewer than the internal evaluation report) in the text to highlight major findings.

Comment: Overall, it appears that the accuracy and the impact of any summary process depends upon the amount of time available for information feedback to the client, what kind of heuristics accompany the data, how the information is "framed", and the form of information display. At this juncture, photos seem to help match the client's initial certitude (and more hidden agendas) with resultant outcomes. In the case of project directors or funding agency program officers who cannot be on-site frequently, photographs helped mitigate the "jaundiced attitude" administrators have about typical evaluation reports. Rhetoric was backed up with visible evidence of process and product behaviors, a phenomenon very rare in education. Along with other program evidence, photos helped to triangulate the findings, including the presence of client's tunnel vision and blind spots (i.e., selective perception). Stakeholders tended to take the evidence "as is", demanded additional evidence less, and appeared to be more open to negative feedback and suggestions for improvement. One part of the explanation appears to be the relative ease in discerning the differential validity of photographs versus other information sources, especially after repeated exposure.

Evaluation experience indicates that this two tiered approach to data sharing is important for both cognitive and political reasons. The data is most likely to be heard and used by clients if they are satisfied with its truth value and utility value. This satisfaction index seems to rise when a client feedback about data accuracy is sought before presenting data of a judgmental nature. In terms of summarizing and reporting findings to the internal clients (current instructor, project director, and future trainers), we found that the use of photos along with narrative text, especially that which was backed up by quotes from participants, helped clarify the findings. It actually appeared to reduce the amount of feedback time required by the clients as well as enhanced the receptivity to results. This seemed to be especially true when the written and verbal feedback included photographs of the clients interacting with participants.

Issues Related to Photo-Interviewing

Our task as instructional researchers and evaluators concerned the issue of how to use photo-interviewing as a tool of inquiry that promoted a shared reality among stakeholders. We sought to use photographic evidence as a tool which made episodes overt among the significant players. Several additional issues should be considered in implementing photo-interviewing. These include cognitive, ethical, and technical issues.

1.Cognitive Issues

The purpose of this section is to explore cognitive dilemmas of photo-interviewing. To seek meaningful patterns of human behavior, one must first start with human perceptions of reality. Specifically, three cognitively complex processes must be examined:

- the perceptual representation of exchanges (between information inputs and feedback messages) by interviewers and interviewees;
- the self-awareness of levels of perceptual representation; and
- the problem-solving skills of the interpreter regarding the data.

Perceptual Representation

Perceptions and judgments of photographs are a function of the interviewer's and interviewee's contexts. A variety of dilemmas assault our senses as we try to make meaning of our world. By making overt the vulnerable representations of both the photo-interviewer and those being interviewed, a synthetic picture of multiple viewpoints and values can be constructed. Through common stimuli (e.g., interviewer questions and photographs), these values can be articulated and help focus the study. In addition, these values can serve as judgment criteria for decision-making. On the other hand, if the operative values are not known, distortion results in terms of what is questioned, how evidence is gathered and how judgments are made. In other words, the value of a photograph varies greatly across viewers. One wonders if the repeated presentations of photographs to interviewees will confirm Sontag's (1977) speculation that repeated exposure to images anesthetizes one's representational capacity. What is the saturation point of the viewer to images? Without a doubt, each photograph has multiple meanings and the dual capacity to make reality manageable as well as opaque; continuous as well as discrete; simple as well as complex; and revealing as well as concealing.

Self Awareness

To come to grips with these operative values and multiple realities, individuals must become more self-aware of their subjectivities. Subjectivity has been defined in the dictionary as "the quality of an investigator that affects the results of observational investigation". If we accept Webster's definition, then it is fairly safe to assert that subjectivity operates throughout the entire photo-interviewing process. Subjectivity has the capacity to filter, skew, shape, block, transform, construe and misconstrue what transpires from the outset of a project to its final report (Peshkin, 1988). If subjective qualities permeate the investigation process at both conscious and unconscious levels, we must search for ways to control the potential bias by making these qualities more overt. One way to cope with subjective bias is to strive for enhanced validity through multiple perceptions. One could conduct many individual photo-interview sessions and aggregate the data OR use group photo-interviewing to permit each perceiver to negotiate his or her subjective reality with others in a tangible and immediate manner. By viewing photos of critical incidents in a sequence with other participants, the learner and investigator alike can transcend the subjective biases resulting from individually dissociated and temporally "frozen" perceptions. The chances of distortion can be further decreased when the data collection method uses more than one sensory channel (such as auditory and visual channels involved with photointerviewing). Finally, the shared feedback of the photos appears to result in significantly better self-awareness than reduced feedback treatments (Harmon and Rohrbaugh, 1987).

Problem-Solving Skills

Photo-interviewing results are highly contingent upon the problem-solving and interpretative skills of the interviewer and interviewees. Photo-interpretation requires information search, and such search is guided by hypotheses which are suggested by semiotic cues. For interviewers and interviewees alike, the skill of generating hypotheses can range from novice to expert. Evidence in general problem-solving research suggests that problem-solving is primarily schema-driven for the expert and search-driven for the novice (Larkin, 1979; Sweller and Levine, 1982; Gilhooly, 1990). This finding has many implications for interpreting interviewee responses and evaluating the quality of interviewer questions. Expert interpreters should have a better memory for new information and a tendency to more clearly perceive the deep structure of the problem (i.e., the photograph to be interpreted). Neophyte viewers focus on key words in the question and surface structure features of the photograph.

In fact, the skills involved in photo-interviewing appear to be very similar to clinical problem-solving. Elstein, Schulman, and Sprafka's results (1978) suggest medical experts generate a few hypotheses within the first few minutes of contact with the patient based on presented symptoms and signs (reasoning forward). These hypotheses are then tested by deducing what other symptoms ought to be present and testing for them (reasoning backwards). Further, it is suggested that experienced physicians develop "illness scripts" containing prototypical information about real patients so that the problem-solving process becomes fairly automatic (Feltovich and Barrows, 1984; Hobus, Boshuizen and Schmidt, 1989). Since evaluation novices lack these real-life experiences, neophytes cannot proceed beyond superfical processing nor develop programmatic "scripts". We believe the photo-interviewing strategy assists in training novice researchers and evaluators to develop viable scripting skills. Photographs combined with verbal data can be overtly manipulated in a variety of contextual and sequential formats, helping them to perceive cues on many levels and reason both backwards and forwards.

2. Ethical Issues

The purpose of this section is to explore ethical dilemmas of photointerviewing. While the ethical implications of interviewing and observation have been studied for many years in biomedicine (Fang & Ellwein, 1990), visual anthropology (Collier & Collier, 1986) and visual sociology (Gold, 1989) , the polemical contexts of educational evaluation and research pose three unique questions.

- How to achieve an accurate representation of events from a research versus an evaluation perspective?
- How can informed consent be truly "informed"?
- How to increase evaluator/research vulnerability along with client/respondent vulnerability?

Accurate representation of events:

In the earlier section on cognitive issues, we discussed the impact of individual perceptions on event portrayal. In this section we will attempt to move from the psychology to the ethics of photo-interviewing. Achieving accurate portrayal requires the synthesis of interviewer and respondent perceptions. Given that researcher and evaluator perceptions are so context bound, accuracy demands that the contexts be representatively sampled in terms of time, purpose, activity/location, method, controls, and ends. Sampling over time and location is important because a program may change from the beginning to the end (Becker, 1978). The purpose of the photo-interviewing and specifically how its use is explained to participants can also impact the accuracy of results. Accuracy demands sampling both private and public (posed and candid) views of participants, which can be perceived as intrusive during the photographic act as well as risky for its ultimate reception when shared during the photo-interviewing session. The ethical risks are compounded when planned and unplanned activities must be captured, especially when it is perceived as "catching someone out". According to Cassell (1982), the "parity of the research relationship is based on the perceived power of the investigator and control of the setting...Persons must always be treated as ends in themselves, never merely as means." Evaluators have much less control of their respondents, which is more representative of a natural setting. Fang and Ellwein (1990) urge consideration of the costs and benefits to all participants in terms of financial, physical, intellectual, psychological and time factors.

Informed consent:

The quest for accuracy is tempered by the potential violation of confidentiality assured in the "informed consent form" signed by participants. Because of both the public and private nature of photographs, it is not possible to conceal the identify of participants nor the setting (Gold, 1989). But while the participant may be knowledgable about the initial concept of the study and initially agreeable, the program changes over time, activities, competence, and trust-building capacity of the photo-interviewers. Gold emphasizes that this issue is compounded by the phenomenon that: a) photo-interviewers and informants are likely to hold different meanings to a picture; and b) photos can be easily gathered without participant awareness. The affect triggered by a picture can vary "profoundly and unpredictably" during the photography as well as after the photos have been developed. Cassell and Wax (1980) advocate substituting the one-shot consent form into a negotiated "covenant" between the observed and the observers. This relates to Gold's coventantal relationship whereby the evaluator and informant develop consensus about the framing of the photograph, the final print, its use and its interpretation. By asking participants to explain their interpretations singly and in small groups, insights can be confirmed about the program's integral contexts, processes, and products. The high degree of interdependence required to cope with this ethical dilemma is much easier to achieve with an evaluation paradigm than an research paradigm which posits little or no control by the subjects. And even then, the evaluator must be very sensitive to each informant's culture and individual reactions. Accurate portrayal requires an evaluator who is not only sensitive and capable of generating meaningful exchanges, he/she must be able to read the participants "visual code" over multiple photos and different levels of visibility (e.g., overt, opaque, covert).

Vulnerability

Who controls the nine phases of photo-interviewing? Where can negotiation play a role? How can the openness of both the observed and the observer be assured? Should a narrative accompany each photograph? And when can vulnerability become a positive condition rather than something to be avoided at all costs? These questions all have vulnerability and control issues in common. Our field studies have shown that the risks of non-negotiation with clients and informants is far greater than the reverse condition. Photo-interviewers who share with informants the selection of critical photographs (candid and public) as well as their use in subsequent reports and presentations lower the chances of misrepresentation, oppression, and harm. Verbal or written narrative which accompanies a photograph further minimizes a photograph being taken out of context. Becker (1986) goes on to add the caveat that both photos and corresponding text should be approved prior to publication or dissemination. Our experiences led us to develop another method for reducing photo-interviewer distortion and oppression--using two photographers to capture an event, using the photo-interviewers and sampling of the target group to select photographs for inclusion in the interview, and interviewing in teams.

3. Technical Considerations.

The purpose of this section is to consider:

- recent developments in photographic technology which may affect photographic data collection and interviewing;
- employing still versus motion stimuli for photo-interviewing.

Recent Developments.

Among the most invigorating aspects of using photo-technologies for data collection and photo-interviewing are the enormous changes going on with the technical capabilities of related media. The dust has not settled on the answers to ways photography will customarily be viewed in the near future. Even so, the handwriting is on the wall for traditional continuous tone images printed on photo-sensitive substrates. This is not an argument for technology for its own sake. Simply, we make comments here because changes are occuring with the processes with which we work. These technology-driven changes, and the long-term reduction in cost per image which will likely accompany them, will encourage greater use of visual media for evaluation and research purposes.

One of the most important recent developments in this regard is the September 1990 announcement by Eastman Kodak of *Photo CD*, a format to be available next year that allows the average home photographer to take up to 100 photos on a special compact disc and inserts into a combination audio CD/Photo CD player and can then be viewed on a television set.

Another rather recent development is the somewhat recent proliferation of "frame grabbing" programs which pull stills or short video segments off of videotape or videodisc. Made possible by innovations such as IBM's M-Motion Video card, stills or video may easily be incorporated into custom software hypermedia programs such as Toolbox on the IBM or Supercard on the Macintosh. The potential for scanning imagery for software programs has been commonplace for some time, but these developments greatly simplify the processes associated with capturing and inserting researcher or evaluator-generated visuals. Other factors which will push the use of photo-related technology are the familiarity with visuals of the "TV generation", the now everyday use of computer technology by all but the most resistant, and the greatly reduced cost of the large amount of computer memory necessary to include graphics in software programs.

It is reasonable to expect that research and evaluation models and methods will adapt to the unprecedented potential accompanying these technological innovations. Microethnographers are already using sophisticated textual analyses to examine speech events recorded in situ and transcribed (Levine, 1990) . Likewise, qualitative researchers and evaluators in education and related social sciences increasingly will use the photo-related data collection and the computer to gain insight into the structure of learning and learning behavior.

Still versus motion

As we have mentioned, it is our preference to gather images for photo-interviewing sessions using a still camera. There are three reasons for this approach. First, a still camera is small, highly-mobile, and forces the data collector to concentrate on "collecting" a single, most valuable image at a time. A data collector is forced by the medium to look for the visual epitome of an occurrence. Second, photographing snippets of information with stills instead of great buckets of data with video frees the evaluator or researcher to work with a more manageable amount of information. Paring down hours of video for use as stimuli in a photo-interviewing session in which even more video will be shot becomes an overpowering, time-consuming task. Third, laying out several photographs on a table or pinning images to a wall during a photo-interviewing session is a relatively natural way to observe diverse imagery. It is fluently understandable to subjects in the way that a still montage is. Subjects invariably home into a single photograph or details of a single photograph. It is the same phenomena that is used to identify a suspect in a crime or increase the power of a microscope to make details clearer. Humans naturally want to dissect information in a visual image. Those who recall David Hemmings' obsession for making the strange familiar in Antonioni' s "Blow up" can imagine the disorientation that might be present in attempting to really dissect a moving image. Few of us not reared on Nintendo games would be capable of it. Perhaps most of us can only truly concentrate on that which is not changing. In effect, we become disoriented by the impossibility of totally concentrating on what is changing.

Conversely, in some cases where time, motion, and sound are important, film or video is irreplacable. Video and film, properly employed can capture a choreography of interpersonal relations in a way that stills usually can not. Artfully shot footage is wonderful for recording the tempo of an event. In essence, the arguments against using moving imagery for photo-interviewing mean little when motion and sound are needed for some reason.

Conclusions and Practical Implications

Several preliminary implications, drawn from issues discussed in this paper, are communicated below.

Given the aims of educational evaluation and research vary from that of the pioneers in photo-interviewing (i.e., biomedicine, sociology and anthropology), the original procedure delineated by Collier needed some modification. The nine step process presented in this paper seems to address the needs of instructional technology more directly.

Photographs can be used as more than data sources. They can also be used for grounding contexts and illustrating concepts across groups of informants and potential "inductees". For evaluators and researchers, the photo-interviewing process creates a common and concrete reference point for subsequent comments and analysis. Wider use of this methodology will certainly be enabled by recent technical developments in photography as well as computer manipulation of qualitative data. One application currently being explored by the authors involves blending photo-interviewing with interactive-video disk technology.

Photo-interviews yielded "richer data" than that obtained from verbal interviewing procedures. One reason is that informants tended to examine images more carefully than the written or spoken word. Further, it appears photo-interviews can generate data particularly helpful for pilot-stage projects where innovation and development rather than maintenance are the priorities.

Photographs can dissolve the alienating or closed verbal authority of the evaluator and researcher, likened to the effect of lecture loving professors who in the process of talking "squelch" or reject interpretations of students who perceive events differently. By providing a common stimulus for all informants, images both solicit significant differences and similarities in individual perception which can be statistically analyzed across groups of perceivers. The methodology provided a means of "getting inside" the program and its context to describe and explain the program and its consequences in terms of participants' realities and meaning systems that oral interviewing did not permit.

To apply this methodology accurately, three cognitively complex processes must be examined: the perceptual representation of exchanges between photo-interviewers and respondents; assessing and enhancing the self-awareness levels of perceptual representation; and systematically building the problem-solving skills of the interpreter regarding potentially overwhelming data.

Photo-interviewing can serve as a research as well as an evaluative strategy given its projective nature. Photographs seem to have a life of their own beyond the evaluation or the project. For example, those photographs which generated negative perceptions (e.g., viewed as inaccurate, disturbing or infuriating by informants) were often indicative of rejection of overt data at odds with their own internal data or belief structures.

Photo-interviewing does not work in all instances. The different contexts of research and evaluation distinctly effect its viability as illustrated by our two episodes. And ethical questions need to be repeatedly raised in terms of: how accurate portrayal of events is achieved; how to make "informed consent" of participants "informed"; and how to reduce exploitation by increasing the interpersonal vulnerability of the observed and observers.

Lastly, enormous technological innovations, including radically different photo materials and hardware, are occuring. Technology-driven changes, the long-term reduction in cost per image, and improved methods of analysis will encourage greater use of visual media for evaluation and research purposes.

In conclusion, although it is unusual to end a paer with a quote, we feltthe following statement by with John Collier was particularly germane.

> I recommend a process of photographic tracking and selective response to 'peaks of information' where and when they happen. This is a flexible approach, one that can be responsive to unpredictable and uncontrollable behavior in social and cultural field settings. It is a process in which selectivity is not an immaculate and mechanical exercise, but a manifestation of human sensitivity, one that rightfully portrays significant research as an accomplishment of human judgement. (Collier, 1979, p. 163).

References

Becker, H.S. (February, 1978). Do photographs tell the truth? Afterimage, 5, 9-13.

Becker, H.S. (1986). Stand in a different place: The view from social science. Paper presented at the annual meeting of the Society of Photographic Education, Baltimore, MD, March 20-23.

Byers, P. (1964). Still photography in the systematic recording of an analysis of behavioral data. Human Organization, 23, 78-84.

Cartier-Bresson, H. (1952). The decisive moment. New York: Simon & Schuster.

Cassell, J. (1982). Harms, benefits, wrongs and rights in fieldwork. In J.E. Sieber (Ed.), The ethics of social research: Fieldwork, regulation, and publication, p. 7-31. New York, NY: Springer-Verlag.

Cassell, J. & Wax, M.L. (1980). Toward a moral science of human beings. Social Problems, 27(3), 259-264.

Collier, J. (1979). Evaluating visual data. In J. Wagner (Ed.), Images of information: Still photography in the Social Sciences. Beverly Hills, CA: Sage Publications.

Collier, J. & Collier, M. (1986). Visual anthropology: Photography as a research method (revised ed.). Albuquerque, NM: University of New Mexico Press.

Converse, J. M. & Presser, S. (1986). Survey questions. Newbury Park, CA: Sage Dick, W., & Carey, L. (1990). The systematic design of instruction. (3nd Ed.). Glenview, IL: Scott, Foresman.

Elstein, A.S., Schulman, L.S. & Sprafka, S.A. (1978). Medical problem solving: an analysis of clinical reasoning. Cambridge, MA: Harvard University Press.

English, F. (1988). The utility of the camera in qualitative inquiry. Educational Researcher, 17 (4), 8-15.

Erickson, F. & Mohatt, G. (1982). Cultural organizational of participant structures in two classrooms of Indian students. In G. Spindler (Eds.), Doing the ethnography of schooling: Educational anthropology in action. New York: Holt, Rinehart and Winston.

Fang, W. L. (1985). The use of photography as a qualitative evaluation technique. Paper presented at the annual meeting of the Evaluation Network and The Evaluation Research Society, Toronto, CA.

Fang, W.L. & Ellwein, M.C. (1990). Photography and ethics in evaluation. Evaluation Review, 14 (1), 100-107.

Feltovich, P.J. & Barrows, H.S. (1984). Issues of generality in medical problem solving. In. H.G. Schmidt and M.L. de Volder (eds), Tutorials in problem-based learning. Assen, The Netherlands: Van Gorcum.

Flanagan, J.C. (1954). The critical incident technique. Psychological Bulletin, 51, 327-358.

Garfinkel, H. (1967). Studies in ethnomethodology. Englewood Cliffs, NJ: Prentice-Hall.

Gilhooly, K.J. (1990). Cognitive psychology and medical diagnosis. Applied Cognitive Psychology, 4, 261-272.

Gold, S.J. (1989). Ethical issues in visual field work. In New technology in sociology: Practical applications in research and work, ed. G. Blank, J.L. McCartney, & E. Brent. New Brunswick: Transaction Books.

Hall, E. (1974). Handbook for proxemic research. Washington, DC: Society for the Anthropology of Visual Communication.

Hobus, P., Boshuizen, H.P. & Schmidt, H.G. (1989). Mental representation of prototypical patients: expert-novice differences. Paper presented at First European Congress of Psychology, Amsterdam, The Netherlands.

Larkin, J.H. (1979). Information processing models and science instructions. In J. Lochhead and J. Clement, (eds), Cognitive process instructions. Philadelphia, PA: Franklin Institute Press.

Levine, H.G. (1990). Models of qualitative data use in the assessment of classroom-based microcomputer education programs. Journal of Educational Computing Research, 6 (4), 461-477.

Mathison, S. (1988). Why triangulate? Educational Researcher, 17 (2), 13-17.

Peshkin, A. (1988). In search of subjectivity--one's own. Educational Researcher, 17(7), 17-21.

Redfield, R. (1960). The little community: Viewpoints for the study of a human whole. Chicago, IL: University of Chicago Press.

Sontag, S. (1977). On photography. New York, NY: Farrar, Straus & Giroux.

Sweller, J. & Levine, M. (1982). Effects of goal-specificity on means-ends analysis and learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 8, 463-474.

Tucker, S., & Dempsey, J. (1990). Photo-Interviewing as a tool of inquiry. Manuscript submitted for publication.

Zemke, R. & Kramlinger, T. (1985). Figuring things out. Reading, MA: Addison-Westley Publishing Company.

Ziller, R.C. (1990). Photographing the self. Newbury Park, CA: Sage Publications.

Appendix A: Section of a photo-interview script which emphasizes critical incidents.

In the next set of questions we are going to lay some pictures of your group working on the table. These photos are numbered. Take a minute to look at them.

If any of them remind you of something or are useful to you in making a point, use them in any way you want. I'll read off the number when you talk about the photograph so that we can tell which photo you were talking about later when we look at the video tapes.

1. Strategy

1P1: Look at these photographs. Do any of these photographs remind you of when you were having trouble with or else when you starting to feel comfortable [using the concentration game or coming up with examples] ? {Pause}

(1a) Think of the time that you got stumped [using the concentration game or coming up with examples]. {Pause} When did this incident happen? What happened?

(1b) What happened to help you figure out how to use this [using the concentration game or coming up with examples]? Or did you use another method? {Pause}

(1c) Think of the time this [using the concentration game or coming up with examples] was the easiest for you to use. {Pause} When did this happen? {Pause}

(1d) Think of the time you felt comfortable in using this [using the concentration game or coming up with examples]. {Pause} When did you see the forest through the trees? What happened? {Pause}

(1e) Think about some strategies you might normally use in learning material of this type? {Pause} Tell me about some occasions when you used your own strategies to learn science material. {Pause}

Title:

Square One Television and Gender

Authors:

Ann De Vaney
Alejandra Elenes

Square One TV is an informal instructional mathematics program broadcast daily by PBS for eight to twelve year olds. Developed two years ago by the Children's Television Workshop (CTW), it aims to change viewers attitudes toward problem solving and to show the usefulness of math in daily life. Although the program was created to help U.S. students "catch up" to their counterparts in other countries, CTW designers decided, at some point in time, not to teach children math, but to change their attitudes. It is clear, then, that the program is affective in nature and only secondarily cognitive.

Bias

Square One TV is broadcast daily for 30 minutes after school and after the screening of *Sesame Street.*. Its format is hybridized. It borrows the structure and conventions of 1980s commercial television programs such as MTV, *Saturday Night Live*, game shows (particularly *Wheel of Fortune*), police dramas and children's cartoons (particulary male superhero shows). Additionally, it parodies *Sesame Street,* 1970s comedies, such as *Laugh In*, and classic detective films of the 1940s (particularly those with Humphrey Bogart).

In a typical *Square One TV* program (#162), 8 to 12-year-old viewers were presented with the following audio or visual sequences at the times indicated.

----45 sec.	The Spanish language is parodied with "Tierra del Freezo" in the visual track.
---47 sec.- 7 min.18 sec.	The Norwegian language and cultural habits are parodied in a sequence about "Vikings."
---8 min. 33 sec.- 11 min. 7 sec.	The audio track of an MTV-like song delivered by a young man includes, *that chick, Maria, ... like an apple she is rotten to the core.* The visual track of this song includes nurses in short revealing white uniforms ministering to the young man by dancing around him.

Approximately one third of a program whose objective is to change student attitudes toward mathematics is either blatantly sexist or racist. Further analysis of *Square One TV* has indicated that the percentage of bias in all programs sampled varies from 20% to 40%.

Project

This paper will describe one stage of a post structural analysis of *Square One TV* focusing on gender issues in the visual text. At first glance, it is difficult to understand how CTW, with twenty years of sensitive children's programming to its credit, could produce such a biased program. *Square One TV* design documents certainly contain no intent to be sexist or racist. Early designers, in fact, are surprised that the program bias exists. Yet, a decision to borrow the structure of 1980s commercial television programs is not a neutral one. Since television programs, like other cultural forms, encode meaning in their structure, they carry with them paradigmatic meaning which has grown up around the use of a particular format. Post structural textual analysis has the power to uncover the path by which sexism and racism were unwittingly enfolded in *Square One TV.*

The post structural reader theory which we are employing to analyze *Square One TV* is reception theory, a German social reader theory practiced by Jauss (1982), Ingarden (1973), Iser (1978) and other textual analysts. (We will also apply this to *Sesame Street.*) It has the uncanny ability to account for transformations and history, an ability that is glaringly lacking in deconstruction. This theory allows us to consider the relationships among viewers, designers and the TV text of *Square One TV.*

The TV text itself is analyzed for codes or syntax patterns emerging from the use of such structural elements as frame, lighting, camera angle, shots, shot transitions and scenes. Codes are traced to their domain of origin, sinceTV codes are seldom original, and examined for paradigmatic meaning. Unlike other post structuralists, Jauss (1982) does not jettison semiotics when investigating the relationship between reader and text. He believes that the text continues to expand the "horizon of expectations" of readers and, conversely, that their readings continually act on the text to change the parameters of its meaning.

Square One TV viewers will first be interviewed and using the results of those interviews, we will survey a larger audience. Since current educational tests to measure affect are based on psychological theories of motivation and/or learning, they ignore the social formation of knowledge and, consequently, the presence of communal readings. (Since CTW producers have declared this program an affective one, it precludes math educators from testing for cognitive gains and excuses CTW from educational accountablity.) Affect should be assessed in some way. Jauss (1982) provides us with a description of modalities of identification, such as associative, admiring, sympathetic, etc. and receptive dispositions, such as transfer, admiration, pity, etc. Our interviews and surveys will be based on Jauss' concepts of reception.

Additionally, *Square One TV* design documents will be examined and, if cooperative, designers interviewed. Relations between and among these three sources will be explored.

Again, this essay will describe one stage of the larger study. It will investigate the program itself, *Square One TV (SOTV)*, for issues relating to gender. Subsequent reports will be written about the relationship between and among the text, the viewers and the sources of production. It is not surprising that commercial television culture leeched into a supposedly benign program designed for 8 to 12-year-olds. It is sad that it has and quite difficult to explain. We believe the analytical tool presented here can trace the process of that cultural path.

PROGRAM DESCRIPTION

The three hallmarks of post modern texts are fractured narratives, fragmented reader subjectivities and dependence, in some way, on high technology. (Nichols, 1988)) After initial viewing, one might call *Square One TV* the quintiessential post modern children's program, since it employs the latest TV technology to create fractured narratives and messages which, in turn, fragment the subject positions of the viewers. To capture the attention of children watching television after school, information about mathematics is incorporated in 1970s and 80s broadcast television formats and presented in segments of approximately 45 seconds to three minutes.

Reading or Viewing

Since television is a communication medium, it is subject to rules governing the sending and receiving of messages. Like any communication medium, it has its own language and grammar. That grammar is, however, renegotiated in daily practice, just as meaning is slowly renegotiated in the use of any language. As the signs and symbols of television are patterned in daily use, codes accrue. Meaning is created when viewers interact with or read television codes which are familiar to them. Codes are usually format specific, such as those specific patterns and conventions which grow up around the production of talk shows, news broadcasts, or sports programs. Viewers becomes familiar with format codes and with the way messages are created within a format. They, consequently, have expectations for understanding, talk shows, news broadcasts or sports programs. A viewer's "horizon of expectations" (Jauss 1982) is pushed by an innovator within one format, such as David Letterman or Roseanne Barr. These performers break the codes or conventions of the current format and begin to redefine it.

Fractured Messages

One post modern technique, particularly visible in architecture, novels and popular media, such as television, is the mixing of forms. Two innovative areas of broadcast television which specialize in mixing formats are commercials and MTV. Within one segment of either of these, the viewer may be offered dramatic narrative, demonstration, comedy skits, animation, documentary or other formats. When forms are mixed numerous things happen to both the message which is being constructed and the viewer who is interpreting that message. If a plot is variously presented in dramatic narrative, documentary and instructional forms, the story is totally fragmented. Such fragmentation of narrative has been used to best effect by some post modern novelists such as Italo Calvino and Peter Handke, but the world of television appears to invite this fragmentation. *Square One TV* has approximately 20 formats which it incorporates and it may include 14 to 16 of them within a 30 minute program. Most of the formats are broadcast television formats or filtered through broadcast television formats. Some of them are summarized in Figure 1 (see Figure 1). Program segments run approximately 45 sec to 3 min. and indeed offer fractured messages. The structure of *SOTV*, furthermore, consists of and only of broadcast television forms. It is as if the designers conceived of television as the whole culture of the child. They use television as one large cultural text which can only reflect itself. It is true that media and communication technologies are so vast and powerful, that our culture is primarily a culture of images or representations. But the producers of *SOTV* took this notion too far when they decided to awaken students' interest in math by reference only to the world of television. These formats incorporate meaning in their structure and whether they are employed in their original programs or adopted as part of *SOTV*, they use the same codes and conventions, the same visual grammar.

Reading Positions

As viewers try to make sense of messages being transmitted by any television program, they rely upon familiar visual codes to interpret or create meaning from what they see. Reader theorists would say that these viewers are the subjects of their own interpretations or meaning. (Tompkins, 1980) They do, however, rely upon the social and cultural meanings of visual codes being employed in these programs. If they care to interpret messages as they are constructed, they are reading "with the program" or accepting the dominant message. If they read against the program or "against the grain," they may reinterpret messages in light of codes which have been established in some subculture such as gays and lesbians (Ellsworth, 1988) or high school students (De Vaney, 1987). Gays reappropriate dominant messages in some elaborate Hollywood musicals, for their own means. High school students read the films of John Hughes differently than adults, because they are privy to and help establish the codes of teenage communication in their everyday lives. But what about young TV viewers watching *SOTV* after school? They are indeed sophisticated television viewers and can switch from format to format with a speed and facility that would astonish most of us, (note the initial popularity of MTV), but while they are adjusting to formats and switching codes, they are probably not reading "against the grain." They are simply making sense of the dominant message as it is presented. This is, most likely, part of the aim of *SOTV* producers, namely, to give audiences short doses of mathematics and hold their attention by switching formats. Viewers are, in other words, offered an invitation (Ellsworth 1988) to accept or buy the dominant message offered in each format.

If viewers are the subjects of their interpretations, they occupy what is called suject positions in relation to each interpretation. Since codes are format specific and messages are constructed by codes, the presentation of many formats in *SOTV* fractures the subjectivity of the viewers. It offers them a vast and, perhaps, confusing range of subject positions. Just what types of invitations are offered to viewers of these vacillating formats? What are the subject positions these children are being asked to occupy? This analysis hopes to answer some of these questions.

Method

As indicated above, the textual analysis of *SOTV* is part of a larger project designed to investigate the way viewers make sense of and, consequently, learn from *Sesame Street* and *SOTV*. The method employed in this larger project is borrowed from a specific reader theory, reception theory, which considers reationships between producers, visual text and viewers/ readers. This essay will report part of the analysis (gender issues) of the visual text. (Further essays, AECT 1990 and AERA1990) will complete the analysis of all three areas. The method used for textual analysis here was as follows.

- Consecutive programs were sampled during two summer weeks.

- After initial viewing, the team formulated focus questions about messages surrounding gender.
- Segments in which these messages appeared were l located and since messages are format specific, formats were identified.
- Segments were reviewed for visual codes within format.
  - syntax patterns within shots were identified
  - syntax patterns between shots were identified
  - syntax patterns in scenes or sequences were identified
- Codes of similar formats on broadcast TV were breifly described.
- Relationships between codes on broadcast TV and *SOTV* were drawn.

GENDER BIAS

SOTV programs incorporated sexism or racism in 20-40 percent of each program, therefore, many segments were available for analysis. We selected segments in which the bias was blatant and some in which it was subtle. The two segments described below were two in which sexism was blatant.

Jonathan Apple Rap; *Rotten to the Core*

Summary. In a psychiatrists' office a young man presents himself to the doctor and explains his "problem." At his workplace, a young girl, Maria, peels apples faster than he can. That fact upsets him, but, secondarily, he cares to learn how to compute Maria's rate of work and his to find the rate per hour at which they could jointly work. In MTV style, the young man sings his problem. The doctor and his crew of short skirted nurses who dance around the young man, solve the problem for him.

Format. The format of this segment (2min .40sec.), or more correctly, music video is ostensibly an MTV format. Opening titles consist of a rotating *Square One TV* logo that is similar in structure to an MTV logo. The syntax patterns are MTV codes during opening and closing sequences. Jump cuts articulate the opening series of tight shots of dancing feet. As in MTV, the geography of the plane is not established during the opening, and the feet, as well as viewer positions, are tossed around in the visual plane. During the first 40 seconds, the camera work is slightly complex with multi-angled tight shots of Jonathan and the doctor juxtaposed with matching cuts. One point of view (POV) shot is included. During this same time period, another locale is established for flashbacks. The scene, to be repeated many times throughout the music video, is a medium long shot of Jonathan's workplace and is encircled by a halo of soft white light (iris-in slightly) to distinguish it as a flashback. A clock rotates across this medium long shot and this is the final visual MTV code that viewers will see until the end of the segment.

The major portion (2 min.) of the music video alternates between scenes in the doctor's office and flashbacks. In both locales the actors are presented as if contained on a stage, most often facing and speaking to the viewer, rather than to one another. Syntax patterns include medium or medium close shots, which are three person shots or four person shots with actors facing the viewers. These are medium angled shots "on the level." The static nature of these codes is quite surprising in an MTV format, particularly since some of the actors regularly dance and sing. The only MTV code that appears to be retained is the duration of the shots; they are short and fast. One might ask if the MTV format was adopted to catch attention and dropped because more serious "subject matter" needed to be communicated. Shots in a medium to medium close range, actors addressing the audience, and shooting "on the level" for the sake of credibility are certainly not codes drawn from MTV, but from instructional film or television, particularly from demonstrations. (We believe these codes were established in World War II training films.) The flashback scenes contain specific clues to this instructional format. Even though flashbacks are a dramatic narrative technique, these are particularly static with no subjective and little objective motion and students are prepared for demonstration. Indeed, all of the superimposition of numbers and math calculations in this music video occur in these flashbacks. Perhaps the designers designated these scenes as the serious instructional ones, therefore, the camera and actors remain static. This switch to instructional formatting within segments is very characteristic of *SOTV* and can be seen in almost every segment.

Content. When Jonathan presents his math problem to the psychiatrist, the doctor construes it as a Freudian problem and tells the young man that his working partner, Maria, has "done a number on his self esteem," because she can peel apples at a rate faster than Jonathan's rate. The doctor informs him that he is

.....consumed with apple envy
..... you feel she's all a feeling
....but she's rotten to the core

Maria's peeling skills, he tells Jonathan, will get him fired. The doctor then describes a dream Jonathan will have in which he will crush

Maria and his boss into applesauce. During this advice there are very tight shots of Jonathan's panicked face. Two nurses, dressed in old fashioned white uniforms which are, incidentally, short and tight, provide a refrain for the doctor's advice. They sing that "problems are tearing him [Jonathan] apart and note that he has a "busted ego" and a "broken id." (Remember all of these actors are addressing the viewer.) Whether these nurses are singing or not, they usually frame the doctor or Jonathan, standing on either side of the man. Suffice it to say, that they also make many gratuitous moves while dancing.

Numerical graphics are both superimposed and edited into the flashback scenes. Graphics shots are approximately 30 seconds in duration and they demonstrate to the viewer the manner in which to calculate the combined rates/hour at which two workers produce.

As the segment ends, Jonathan, in a 30 sec. sequence, ignores the Freudian interpretation of his problem and indicates that it is only a math problem. The music video closes, as does similarly on MTV, with a banner in the upper lefthand corner which contains the singer and song title and a similar banner in the upper righthand corner with the word "Exclusive" on it. The banner in the upper lefthand corner reads, Jonathan Apple Rap, *Rotten to the Core.*

Commentary. The purpose of this commentary is to question the visual text itself and hypothesize about the formation of viewer interpretations. (*SOTV* viewers and designers will be polled as this research continues.) Visual structural codes of MTV were employed in opening and closing titles and opening sequences. MTV verbal codes of rap, song and dance were employed throughout. Since codes are format specific, student viewers were prepared to be entertained and were probably delighted by the fractured geography of the plane and the rap. Only 40 seconds into this segment, the format changes to an instructional one, although rap is retained throughout. Children are sophisticated television viewers, a fact which CTW designers appreciated, and the viewers tacitly understand this shift. Video demonstration is something with which they are quite familiar and the mixing of instruction and rap probably delighted them and helped maintain their attention. They do not have trouble reading mixed formats.

But what was actually taught during the instructional sequences of this music video? If the intention of the designers is to interest students in math, then the arithmatic part of the segment was important. Yet these shots were only approximately 30 secondslong, hardly long enough for the adults whom I consulted to calculate the problem. Whether intentional or not, what was transmitted during the instructional sequences of this music video were social and cultural messages. The rap constructs a familiar discourse about women. "That chick, Maria," it tells the children is, "rotten to the core." Maria's offense is that she peels apples faster than Jonathan. In fact, the title of the music video the student is informed in a closing shot is *Rotten to the Core*. In an instructional format a closing shot helps to summarize and reinforce the message of the lesson. Jonathan, additionally, is told that Maria has "done a number" on his self esteem and that he is filled with apple envy. Not only does the music video reinforce sexist notions of women with which these viewers are familiar, but introduces them to new sexist Freudian concepts. To reinforce this discourse, nurses are shown framing a man and are presented in a toned down version of MTV female background singers.

The construction of the whole music video is intricate, because the designers, in this age of raised consciences, paid tribute to racial equality by casting black actors as the doctor and one nurse. Additionally, they displayed their own knowledge of gender issues by allowing Jonathan to ignore, but not reject, the Freudian interpretation of his problem. Since the designers now know better, they may feel that knowledge allows them to joke about sexism. But then, in the form of Jonathan, they indicate that they were only joking. This ironic distancing of author or producer from their subject matter, with an expectation that enlightened viewers will also distance themselves, is a practice emerging in post modern literature and pop culture. Educators, however, must seriously question why students viewing a 2min. 40sec. supposedly mathematical music video were reintroduced to old concepts and taught new concepts about sexism?

The Coquette and the Quarter

Summary. In a dramatic skit which mixes the forms of dramatic narrative and advertisements, a young boy of 12 or 13 drops a quarter down a sidewalk grate. A young woman of the same age walks by and he bemoans the fact that he cannot figure out how much he dropped down the grate. He can neither see nor retrieve the coin. She tells him to subtract the amount he now has in his pocket from the amount he had prior to dropping the coin. He figures he has lost a quarter and she lends him one, so he may rent a film. They walk off together.

Format. This dramatic skit is 1min. 30 sec. in length and mixes codes from dramatic narrative and advertisements uniformally throughout. The segment consists of one scene which is articulated primarily by jump cuts. Shots move through a range of very tight to tight, with the exception of the last medium shot in which the pair walk off together. The scene opens with a tight shot of the boy's shoes, jumps to a tight shot of his panicked face and then to a tight shot of pink pumps beneath a disembodied pair of tight jean legs. As the scene progresses, the viewer is forced to put these body parts together, feet, faces, eyes, lips, the girl's bottom and by doing so, they employ the reading codes of MTV. About 45 sec. into the scene, dramatic narrative dialogue codes are used to present the conversation between the actors. Even though the dialogue shots are POV, over the shoulder shots, the camera maintains its tight or very tight distance. The dialogue codes are not neutral here, but are layered with Hollywood film romance codes. The dialogue is presented as one between lovers. The camera lingers on the girl's adoring eyes or the boy's moonstruck eyes, presents them very close together and frames one scene to appear as if the two were kissing.

Content. The content of this sequence is quite simple. A boy drops a quarter down a grate and is upset by the fact that he cannot retrieve it, nor does he know how much he has lost. A young friend comes along and solves his problem by teaching him how to subtract and lending him a quarter, so he can rent the film he had planned to rent. It appears straight forward, but it is doubtful that subtraction is what is learned by children viewers. The audio and visual tracks are hardly related. An entirely different story is being told in the visual track.

Commentary. It is unimaginable that CTW designers could combine MTV codes with romantic Hollywood film codes and not know what they were doing. Children viewers receive very tight shots of body parts, the girl's pink pumps with tight jeans, her mascarad eyes, her lipstick-red lips and her bottom. The boy is on screen less than she and fares only slightly better. His feet, face and "moony" eyes are shot. Again, the tight shots articulated with jump cuts probably delight the viewers and engage them, as they do in MTV, in establishing themselves in the geography of the plane and piecing together the message. The message involves the sexualization of children. Codes of kiddie porn are also evident in this segment. Children are heavily made up; certain body parts are fragmented and presented as sexual fetishes. What imbues the codes of advertisement and kiddie porn with power, however, is the employment of dramatic narrative codes. This is a story with a plot; boy meets girl; they "fall for" one another and walk off together. It is important to remember that viewers read dramatic narrative most often by identifying with the hero or heroine. Hollywood directors knew that when they established the dialogue codes of POV, over the shoulder shots. A point of view shot places the viewer in the shoes of the speaker. Very surprisingly, this is one *SOTV* segment in which no instructional format appears. Neither character ever addresses the audience, but even when dramatic narrative is invoked in other *SOTV* sequences, the instructional format consistently breaks in. It is difficult to imagine that the designers thought they were teaching subtraction in this segment, or even interesting viewers in subtraction. What is taught is how a 12 or 13 year-old-boy and girl should act, dress and speak in a sexual fashion. Again, the scene involves young viewers in a familiar discourse about gender relations. A young girl should dress and make up in a certain manner. Pumps should be worn with tight jeans. Mascara and lipstick are required and a bow in the hair is helpful. Her stance and looks, when talking to a boy, should be seductive. The poor boy, of course, is powerless in this situation. Certainly United States viewers from eight to 12 are familiar with this discourse, but the clothes and "the look" give them an updated version of the discourse, and the inclusion of children their own age, offers them role models. We must question who made the decision to include 12 or 13-year-old actors and 8 to 12-year-old viewers in this scene with codes from TV ads, Hollywood romance films and kiddie porn.

Patterns

This essay has only been able to address a fraction of the sexism inherent in *SOTV*. The larger study also examines racism and explores designer intentions and viewer interpretations.

Although this essay has only presented an overview of *SOTV* and an indepth analyses of two segments, we believe both the overview and the analyses are representative of the program as a whole. We have found bias in 20% - 40% of all programs sampled. To date, we have found a pattern in these biased segments, the characteristics of which have been mentioned in the analysis here. When sexist or racist materials are presented in these programs, designers most often distance themselves from, or excuse the material in various ways. Racist material is always hidden beneath the obvious inclusion of blacks in these segments. One must infer that the designers are partially color blind. Blacks are never parodied, but other people with accents are, such as Hispanics, Arabs and Italians. Sexist segments, however, are presented in one of two ways. The inclusion of blacks is sometimes used as a foil for sexism, as in *Jonathan Apple Rap*, but sometimes sexism is blatantly presented with no excuses. These two approaches are clearly present on current day broadcast television as well. Some detective shows, police programs and MTV are blatantly sexist. Yet, some sitcoms and ads use ironic distancing to present sexist materials. (Recall the current shampoo ad in which a sultry female voice tells the viewer that

"All chicks like to be soft," while the camera focuses on actual yellow chicks. Later the camera does switch to the beautiful speaker.)

Late 80s performers are allowing themselves to be sexist or racist, perhaps because they know better and expect that their audience knows better. They allow themselves to be shocking. Even Johnny Carson,(Pareles, 1989) while recently playing one of his characters,"Clem," made such racist remarks that the studio audience gasped. He had distanced himself from his material by letting the bumpkin Clem deliver the lines. It is as if enlightened performers and audiences can now laugh at bigotry, because they are no longer bigotted. Other sociological reasons have been suggested for this return to bigotry, (Pareles, 1989) but ironic distancing is a major characteristic of the post modern age and television is the expert producer of the post modern visual text. We believe that the designers of *SOTV* recognized this. They also recognized that they were dealing with children whose primary cultural medium is television.

Children may be living in a post modern age and may be sophisticated viewers of television. But, their sophistication resides in their ability to shift format codes and decode visual information quickly. They are hardly ironic distancers and the pace of *SOTV* almost guarantees that they only have time to process the dominant message. They engage the facts of what they see. And even if they were ironic distancers, such distancing from pattently sexist and racist discoursers works as an ethical sleight of hand for the practice of sexism and racism.

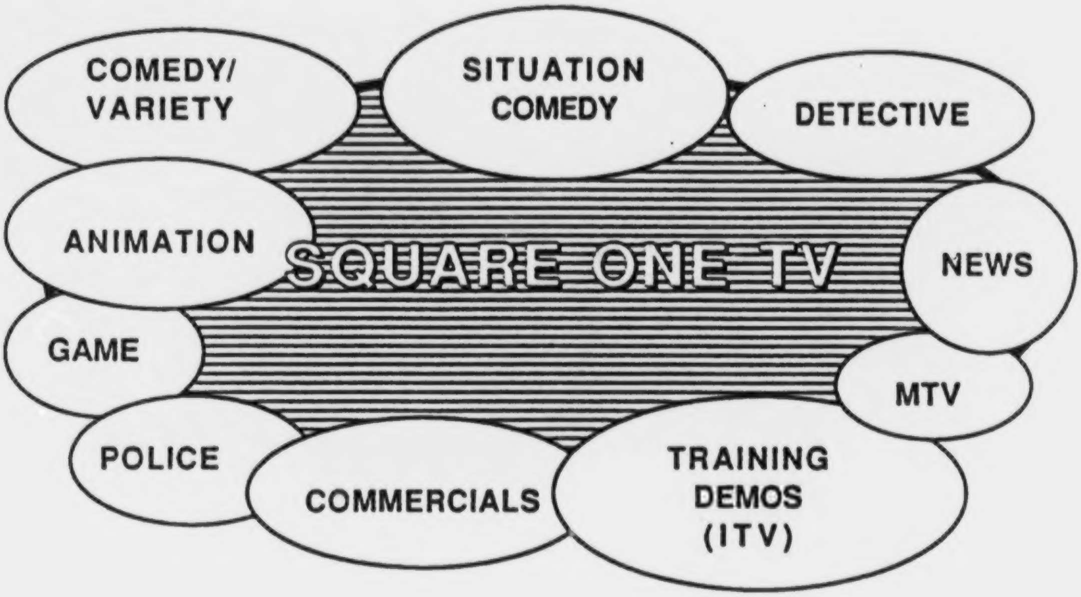

COMEDY/
VARIETY
SITUATION
COMEDY
DETECTIVE
ANIMATION
SQUARE ONE TV
NEWS
GAME
MTV
POLICE
COMMERCIALS
TRAINING
DEMOS
(ITV)

FIGURE 1: Overlapping codes of SOTV and Broadcast Television Formats

REFERENCES

De Vaney, A (1987). The world of John Hughes, *Framework 1*, 1, 16-26.

Ellsworth, E. (1988). Media interpretation as a social and political act, *Journal of Visual Literacy* 8, 27-38.

Ingarden, R. (1973). *The cognition of the literary work of art.*. tr. R. A. Crowley and K. R. Olsen, Evanston, Illinois; Northwestern University Press.

Iser, W. (1978). *The act of reading.* Baltimore; The Johns Hopkins University Press.

Jauss, H. R. (1982). *Towards an aesthetic of reception.* tr. T. Bahti, Minneapolis; University of Minnesota Press.

Nichols, B. (Winter 1988). The work of culture in the age of cybernetic systems, *Screen, 29,* 1, 22-46.

Pareles, J (Sept. 10, 1989). There's a new sound in pop music: bigotry, *The New York Times, Arts and Leisure,* pp.1, 32-33.

Tompkins, J.P. (1980). The reader in history, *Reader response criticism* ed. by J.P. Tompkins, Baltimore; The Johns Hopkins Press, 201-232.

Acknowledgement. The authors of this essay wish to acknowledge the invaluable assistance of Tonja Meyers, assistant editor of the Journal of Visual Literacy, in transcribing data from the television programs and preparing this paper.

Title:

A Comparison of an Expert Systems Approach to Computerized Adaptive Testing and an Item Response Theory Model

Author:

Theodore W. Frick

ABSTRACT

Expert systems can be used to aid decision making. A computerized adaptive test is one kind of expert system, though not commonly recognized as such. A new approach, termed EXSPRT, was devised that combines expert systems reasoning and sequential probability ratio test stopping rules. Two versions of EXSPRT were developed, one with random selection of items (EXSPRT-R) and one with intelligent selection (EXSPRT-I). Two empirical studies were conducted in which these two new methods were compared to the traditional SPRT and to an adaptive mastery testing (AMT) approach based on item response theory (IRT). The EXSPRT-I tended to be more efficient than the AMT, EXSPRT-R and SPRT models in terms of average test lengths. Although further research is needed, the EXSPRT-I initially appears to be a strong alternative to both IRT- and SPRT-based adaptive testing when categorical decisions about examinees are desired. The EXSPRT-I is clearly less complex than IRT, both conceptually and mathematically. It also appears to require many fewer examinees to establish empirically a rule base when compared to the large numbers required to estimate parameters for item response functions in the IRT model[1].

THEORETICAL ISSUES

An Overview of Expert Systems

One of the more practical results from extant research in artificial intelligence is the application of expert systems reasoning to aid in decision making or problem solving. Expert systems have been developed, for example, to help physicians identify types of bacterial infections, to aid investor decisions on buying and selling stock, for aid in assembling components of computer systems, for making decisions about where to drill for oil, for assisting underwriters in making insurance policies, and for diagnosing causes of equipment failures to help repairpersons (cf., [1]).

An expert system consists of a set of production rules or frames, often called a 'knowledge base'. The name, 'expert system', was coined because a knowledge base is typically constructed by interviewing one or more experts in some domain of knowledge. An attempt is made to capture their reasoning processes, when they solve problems in that knowledge domain, in the form of "If..., then..." rules. For example, in MYCIN, a famous early expert system for diagnosing bacterial infections, one of the rules is:

IF 1) the gram stain of the organism is negative, and
2) the morphology of the organism is rod, and
3) the aerobicity of the organism is anaerobic,

---

[1] This study was supported in part by a grant from the Proffitt Foundation, School of Education, Indiana University. Hing-Kwan Luk adapted and extended considerably the author's computer code for conducting the test re-enactments with the four methodologies. Luk also assisted with statistical analyses. Thomas Plew must be acknowledged for his significant contribution of an initial method of intelligent item selection for the EXSPRT, subsequently refined by the author. Had Tom not asked the questions he did in a doctoral seminar taught by the author in 1987, then EXSPRT might not exist today.

THEN there is suggestive evidence (.7) that the identity of the organism is Bacteroides. ([2], p. 34)

This particular rule is one of over 400 such rules that comprise the MYCIN knowledge base. A computer program, called an 'inference engine', uses this rule set as data to help physicians identify unknown bacteria. The program makes inferences by using both the rule set and specific answers to questions it asks the physician about properties of the current situation (e.g., patient symptoms, white blood cell count, and other lab test results). MYCIN has been shown to be more accurate in its identifications of bacteria than typical practicing physicians, particularly in identifying those bacteria which are rarely observed.

Expert systems are not usually viewed as replacements for human decision makers, but as aids or tools for such persons. Expert systems obviously cannot perform in areas not covered by the knowledge base. Furthermore, decisions reached by expert systems can be no better than the accuracy of the knowledge or rules that comprise the database.

In education and training, expert systems principles have been applied mostly in intelligent tutoring systems ([3], [4]). As an example, GUIDON was later developed from MYCIN in an attempt to teach physicians how to identify different kinds of bacteria [5].

Similarities between Expert Systems and Adaptive Tests

One efficient and empirically validated approach to computerized adaptive testing (CAT) is based on item response theory (e.g., [6]). An adaptive test is no longer than necessary to obtain a satisfactory estimate of an examinee's ability, and items are selected which are close to his or her estimated ability level. For example, if a person misses a question, a somewhat easier question is next asked. On the other hand, if a question is answered correctly, then a slightly more difficult question is subsequently selected. A computerized adaptive test does not waste time administering questions that are too hard or too easy for a particular individual. Adaptive tests tend to be shorter than conventional fixed-length tests and the results are as reliable if not more so (e.g., [6]).

Expert systems and adaptive computer-based tests have many properties in common:

1. The rule base for a CAT is a set of item characteristic curves estimated from prior test administrations. That is, each item characteristic curve is a compact way of saying, "If the examinee ability level is *X*, and item *Y* is asked, then the probability of a correct response is predicted to be *Z*."

2. Both expert systems and CATs use inference engines that are often Bayesian or Bayesian-like. Even if rules do not have probabilities (or confidence factors) associated with them, they can still be treated as a special Bayesian case where associated probabilities are either one or zero (cf., [7]).

3. The goal of an expert system is to choose from a number of alternatives (e.g., causes of equipment failure) using the rule base and answers to questions it selects and asks of a particular user. The goal of a CAT is to estimate an examinee's achievement or ability level with enough precision to make a decision such as pass/fail or a grade classification using a rule base of item characteristic curves and answers to questions it selects and gives to examinees.

4. An expert system selects which questions it asks by using forward or backward chaining and the rule base. A CAT can select questions on the basis of the amount of information they provide, depending on the ability of an examinee. For example, Weiss and Kingsbury use a maximum information search and selection (MISS) procedure [6].

Thus, although not widely recognized at this time, an adaptive testing system is one type of an expert system. The author realized this when developing computer code for an expert system, having already developed code for Bayesian decision methodologies and a computer-based testing system.

On hindsight, expert systems and adaptive tests have much in common. Yet in the research literature it appears that these two threads of development have been almost entirely independent. One camp has grown out of an artificial intelligence movement and the other from a psychological testing and measurement perspective. A recent computer search of numerous bibliographic databases only turned up thirteen articles where the terms, 'expert systems' or 'artificial intelligence' and 'adaptive' or 'computer' and 'testing' or 'test' were used as descriptors. The two camps not only use different language to describe their activities, but also tend to publish in different journals and attend different conferences.

THE DEVELOPMENT OF *EXSPRT*

A problem with the IRT-based approach to adaptive testing faced by many practitioners, however, is that a relatively large number of examinees must be tested in advance in order to estimate accurately item parameters of difficulty, discrimination, and lower asymptotes (200 to 1000 depending on the model used and the number of items in a pool). Furthermore, proponents of the Rasch model (one-parameter IRT model) have indicated that there is no valid way of estimating item discrimination and lower asymptotes for the two- and three-parameter models without imposing arbitrary constraints (cf., [8]).

The author has previously investigated the predictive validity of the sequential probability ratio test (SPRT) for making mastery decisions, where the lengths of tests were adapted according to student performance ([9]). Mastery decisions reached with the SPRT, when used conservatively, agreed highly with those based on total test results. Nonetheless, the SPRT does not explicitly take into account variability in item difficulty, discrimination or chances of guessing as does the three-parameter IRT model. Moreover, items are selected randomly in the SPRT, rather than on the basis of their characteristics and estimated examinee ability or achievement level as in the MISS procedure.

Is there some middle ground between the relatively simplistic SPRT decision model and the relatively sophisticated IRT-based approach? When considering the problem from an expert systems perspective, a solution became apparent. Instead of considering a continuum of alternatives, as is the case in IRT-based CAT, it was hypothesized that if the goal of an adaptive testing system is to choose between a few discrete alternatives (e.g., mastery or nonmastery; grades of A, B, C, etc.), then it should be possible to develop a satisfactory rule base from a smaller sample of examinee test data--compared to the IRT model.

Development of the Rule Base for a Given Test Item Pool

Assume that we have developed a pool of test items which match a particular instructional objective and that our goal is to choose between two alternatives for any given student: mastery or nonmastery of the objective. For each item *i* in the pool we create four rules:

Rule i.1: If the examinee is a *master* and item *i* is selected, then the probability of a correct response is $P(C_i \mid M)$.

Rule i.2: If the examine is a *master* and item *i* is selected, then the probability of an incorrect response is $P(\neg C_i \mid M)$.

Rule i.3: If the examinee is a *nonmaster* and item *i* is selected, then the probability of a correct response is $P(C_i \mid N)$.

Rule i.4: If the examine is a *nonmaster* and item *i* is selected, then the probability of an incorrect response is $P(\neg C_i \mid N)$.

Notice that these rules are essentially in the same if-then form as the sample rule from the MYCIN expert system at the beginning of this article. In MYCIN the rules were developed on the basis expert knowledge on the co-occurrence of various kinds of gram stain, morphology and aerobicity and the incidence of each kind of bacterium. In our case, we will rely on empirical data collected on the test items with a representative sample of examinees who are characterized by the discrete categories from among which our expert system will later attempt to choose.

In the dichotomous case, the estimates of probabilities of correct responses to items by masters and nonmasters are determined as follows[2]:

1. Give the pool of test items to a representative group of examinees, about half of whom are expected to be masters and half nonmasters--i.e., for whom you expect a wide range of scores on the test.

2. Choose a mastery cut-off score (e.g., .85).

3. Divide the original group into a mastery group and nonmastery group based on their total test scores and the mastery cut-off.

4. For each item in the mastery group, estimate the probabilities of correct and incorrect responses by the following formulas (see [10])[3]:

$$P(C_i \mid M) = (\#r_{im} + 1)/(\#r_{im} + \#w_{im} + 2) \qquad \{1.1\}$$

$$P(\neg C_i \mid M) = 1 - P(C_i \mid M) \qquad \{1.2\}$$

---

[2] Note that this reasoning can be extended to more than two categories, such as letter grade designations.

[3] Note that the estimates of these probabilities of correct responses to items by masters and nonmasters will never be one or zero. This means that, in the EXSPRT Bayesian updating process during the administration of a test to an examinee, the probabilities of the mastery and nonmastery alternatives will never be zero or one, though these extremes may be closely approached.

where $\#r_{im}$ = number of persons in the mastery group who answered the item correctly;

and $\#w_{im}$ = number of persons in the mastery group who missed the item.

5. Do likewise for the nonmastery group for each item:

$$P(C_i \mid N) = (\#r_{in} + 1)/(\#r_{in} + \#w_{in} + 2) \tag{1.3}$$

$$P(\neg C_i \mid N) = 1 - P(C_i \mid N) \tag{1.4}$$

Method of Making Inferences in the EXSPRT

The reasoning procedure employed in our expert systems approach is Bayesian, with the addition of stopping rules from the sequential probability ratio test (SPRT) (cf., [9, 10, 11]). After each observation (i.e., administration of a test item), a likelihood ratio is computed:

$$LR = \frac{P_{om} \prod_{i=1}^{n} P(C_i|M)^s \, [\, 1-P(C_i|M)\,]^f}{P_{on} \prod_{i=1}^{n} P(C_i|N)^s \, [\, 1-P(C_i|N)\,]^f} \tag{2}$$

where

P_{om} = prior probability that the examinee is a master,

P_{on} = prior probability that the examinee is a nonmaster[4]

and $s = 1, f = 0$ if item i is answered correctly,

or $s = 0, f = 1$ if item i is answered incorrectly.

The three stopping rules are:

If $LR \geq (1 - \beta) \div \alpha$, then stop asking questions and choose mastery. {3.1}

If $LR \leq \beta \div (1 - \alpha)$, then stop asking questions and choose nonmastery. {3.2}

Else ask another question, update *LR*, and reiterate rules 3.1 to 3.3. {3.3}

---

[4] Note that if the prior probabilities of mastery and non-mastery are equal, then they drop out of the formula for the likelihood ratio.

Alpha and beta are Type I and II decision errors, respectively. Alpha is the probability of choosing mastery when the nonmastery alternative is actually true. Beta is the probability of choosing nonmastery when the mastery alternative is true.

For numerical examples of this Bayesian reasoning process, the reader is referred to [12]. Also note that expert systems can use reasoning procedures other than Bayesian, such as the Dempster-Shafer rules. See [7] for further examples of Bayesian reasoning in expert systems.

Random Item Selection: EXSPRT-R

When the EXSPRT was initially conceived, it was seen as an extension of the Bayesian approach to the SPRT, as noted in [9], using empirically derived data for estimating the probabilities of correct responses by masters and nonmasters to each test item rather than *average* probabilities across all items. In this initial approach to the EXSPRT, items were selected randomly without replacement, and it was assumed that observations were independent in order to multiply the conditional probabilities to form the likelihood ratio. This version of the EXSPRT is referred to as EXSPRT-R, in contrast an intelligent item selection procedure discussed below.

A reader who is knowledgeable about expert systems may note that the method of choosing questions in the EXSPRT approach to adaptive testing *departs* from the typical method of forward or backward chaining used in many expert systems. One might argue that the EXSPRT approach is not really expert systems' reasoning because forward or backward chaining is apparently not used to determine which question to ask next in order to reach, or move closer to, a conclusion (i.e., choose from a number of discrete alternative goals).

On the other hand, one who understands expert systems deeply, having written computer code to carry out the reasoning, will conclude that the EXSPRT is doing essentially the same thing as occurs in either backward or forward chaining. The rules in the EXSPRT are all of the same form, however, in contrast to rule bases in expert systems where a variable's value in a consequent condition in one rule is also part of an antecedent condition in another rule, etc. (cf. [7]). One could say that instead of just asking a particular question once, and then immediately making an inference based on the answer to that question--as is usually the case in most expert systems--in the EXSPRT many questions of the same kind are asked until an inference can be made confidently. In the EXSPRT-R questions are chosen at random in order to comprise a representative sample from the item pool.

Intelligent Item Selection: EXSPRT-I

Thomas Plew was not satisfied with EXSPRT-R, since it did not use information about test items in the selection process [13]. This stimulated the joint development with the present author of an item selection procedure that is modeled after basic principles used by Weiss and Kingsbury in the MISS (maximum information search and selection) procedure. Though the principles are comparable, the mathematical approaches are quite different.

In the EXSPRT-I (i.e., with "intelligent" item selection), the reasoning is as follows:

Item discrimination. If we are trying to choose between mastery or nonmastery alternatives, then an item is more discriminating when the difference between probabilities of correct responses by masters and nonmasters is greater. For example, if the probabilities of a correct response to item #5 are .90 for masters and .25 for nonmasters, then item #5 is very discriminating (difference = .65). On the other hand, if the probability of a correct response to item #53 is .85 for masters and .75 for nonmasters, then this item is much less discriminating (difference = .10). Or if the probability of a correct response to item #12 is .60 for masters and .80 for nonmasters, then such an item is negatively discriminating (difference = -.20).

Thus, the discrimination index for item *i* is defined:

$$D_i = P(C_i \mid M) - P(C_i \mid N) \qquad \{4\}$$

Item/examinee incompatibility. Not only do we want to select highly discriminating items, but also we want to select items that are matched to an examinee's estimated achievement or ability level. In theory, we gain little additional information by administering items which are very easy or very hard for a given individual. Better items would be those which a person has a 50/50 chance of answering correctly--i.e., which are very close to her or his achievement level. For example, if an examinee's achievement level is estimated to be .80 (on a scale from zero to one), then a good item would be one that was answered *incorrectly* by 80 percent of the examinees in the item parameter estimation sample [$P(C_i) = .20$ for masters and nonmasters combined].

Thus, the item/examinee incompatibility index is defined for each item:

$$I_{ij} = abs\{(1 - P(C_i)) - E(\Phi_j)\} \qquad \{5\}$$

where $$E(\Phi_j) = (\#r_j + 1)/(\#r_j + \#w_j + 2) \qquad \{6\}$$

and $$P(C_i) = (\#r_i + 1)/(\#r_i + \#w_i + 2) \qquad \{7\}$$

Note that $\#r_j$ and $\#w_j$ are the numbers of questions answered correctly and incorrectly, respectively, thus far in the test by the current examinee. Note also that the estimate of $P(C_i)$ is based on the *total* number of persons in the parameter estimation sample for item *i*, irrespective of mastery status. Thus, $\#r_i$ is the number of persons who answered item *i* correctly and $\#w_i$ is the number who answered it incorrectly. Finally, note that the item/examinee incompatibility index is based on the absolute value of the difference between the estimate of the probability of an *incorrect* response to the item and the estimate of the current examinee's achievement level (proportion correct metric).

Item utility. As a test proceeds, item utilities are re-calculated for all items remaining in the pool, in order to select and administer a new one that now has the most utility for an examinee:

$$U_{ij} = D_i/(I_{ij} + \delta) \qquad \{8\}$$

where δ = some arbitrary small constant (e.g., .0000001), to prevent division by zero in case $I_{ij} = 0$.[5]

Thus, each utility value is simply the ratio of the discrimination of item *i* and its incompatibility with person *j*'s achievement level. The item that is selected next in the EXSPRT-I (intelligent selection) is the remaining one with the greatest utility at that point for that particular examinee. This means that the item selected next is the one which *discriminates best* between masters and nonmasters *and* which is *least incompatible* with the current estimate of that examinee's achievement level. Note that item utilities change during a test, depending on an examinee's performance which affects the estimate of his/her achievement level in the item/examinee incompatibility index. In effect, the EXSPRT-I is comparable to the two-parameter item response theory model (IRT--see [14]) in that both item discrimination and item difficulty are considered in the item selection process.

Unanswered Questions

Since the EXSPRT-R and EXSPRT-I are new approaches to computerized adaptive testing, two empirical studies were conducted to compare these approaches to extant IRT-based adaptive mastery testing and SPRT approaches (cf., [6], [9], [14]). Of major concern was the accuracy with which each adaptive model could predict decisions based on total test scores. Does each adaptive method make mastery and nonmastery decisions with no more errors than would be expected by *a priori* error rates? Second, how efficient is each adaptive method in terms of average test lengths for mastery and nonmastery decisions? Are any of the methods more efficient than others?

FIRST STUDY

Digital Authoring Language Test

A computer-based test on the structure and syntax of the Digital Authoring Language was constructed, consisting of 97 items, and referred to as the DAL test. This test was comprised of multiple-choice, binary-choice, and short-answer questions. The test was highly reliable (Cronbach α = .98). The DAL test was also very long, usually taking between 60 and 90 minutes to complete, and it was very difficult for most examinees (mean score = 63.2 percent correct, S.D. = 24.6).

Examinees

The persons who took the DAL test were mostly either current or former graduate students in a course on computer-assisted instruction taught by the author. Those students who were currently enrolled at the time took the DAL test twice, once about mid-way through the course when they had some knowledge of DAL--which they were required to learn for developing CAI programs--and once near the end of the course when they were expected to be fairly proficient in DAL. The remainder of the examinees took the DAL test once. Since the test was long and difficult, no one was asked to take the test who did not have some knowledge of DAL or other authoring languages.

---

[5] Alternatively, δ_i could be considered as some kind of "guessing" factor for the item. However, this will not be considered in the present paper.

Test Administration

The DAL test was individually administered by the Indiana Testing System [15]. As an examinee sat at a computer terminal, items were selected at random without replacement from the total item pool until all items were administered. Students were not allowed to change previous answers to questions, nor was feedback given during the test. Upon completion of test, complete data records were stored in a database, including the actual sequence in which items were randomly administered to a student, response time, literal response to each item, and the item scoring (correct or incorrect). Examinees were informed of their total test scores at the end of the test. There were a total of 53 administrations of the DAL test in the first study.

Experimental Methods

The basic procedure was to re-enact each test, using actual examinee responses in the database, for each of the four adaptive methodologies: 1) IRT-based adaptive mastery testing (AMT--with maximum information search and selection [MISS]), 2) sequential probability ratio test (SPRT), 3) EXSPRT-R (random selection of items), and 4) EXSPRT-I (intelligent selection of items--see above descriptions).

Item parameter estimation. Two random samples of examinees were used to estimate item parameters ($n = 25$ and $n = 50$), the latter containing the former. This was done to see if increasing the sample size used for parameter estimation would result in fewer decision errors in the four methods. Due to the relatively small sample sizes, the one-parameter AMT model was used--i.e., only b_i estimates were obtained for the two samples using program BICAL [16]. For the EXSPRT-R and EXSPRT-I, the rule base for each parameter estimation sample was constructed using formulas {1.1}, {1.2}, {1.3} and {1.4}. The mastery cut-off was set at 72.5 percent, half way between the established .85 mastery level and .60 nonmastery level used in an earlier study of the SPRT only [9]. In the current study, however, the mastery and nonmastery levels for the SPRT were established *empirically* from the .725 cut-off and the two parameter estimation samples. The mean proportion correct for masters was used as the mastery level and the mean proportion correct for nonmasters was used as the nonmastery level in each sample. In effect, the SPRT was treated just like the EXSPRT-R, except that the rule quadruplets for all items were the same in the SPRT, based on the sample means for masters and nonmasters, respectively.

Test re-enactments. Once the parameter estimation samples were chosen, then two doctoral assistants independently wrote computer programs in two different languages (Pascal and DAL) to construct the rule bases for the EXSPRT, and to carry out the four different adaptive testing methods on the same 53 sets of test administrations. This was done to reduce the possibility of error in coding these rather complex methodologies, especially the AMT model. When results did not agree, as was occasionally the case, this helped to identify and ameliorate errors in coding. The one difference that was not correctable was traced to the precision of arithmetic in DAL and Pascal on a VAX minicomputer.

It was discovered that on occasion the MISS procedure in the two programs would begin to select different items in the AMT model after 15 to 20 items had been retroactively "administered" to an examinee. This occurred because the updating of the estimate of Θ and its variance, and in turn the item information estimates for that Θ estimate, would tend to differ very slightly in the two code versions as a test progressed. Consequently, the MISS procedure would occasionally pick a different item in the two different versions when estimates of item information were very close for two or more items remaining in the pool. From that point on in a test, different item sequences were observed. The average AMT test length in the DAL version tended to be about one item shorter, compared to the Pascal version, but the decisions reached were the same with one exception.

These discrepancies do point out a problem inherent in the IRT-based approach, which contains numerous multiplications, divisions, and exponentials (see [14], formulas {9) to {25}). Very small errors due to rounding or differences in precision of arithmetic can magnify themselves rather quickly. This problem was not observed with the EXSPRT-I, EXSPRT-R, or SPRT--other than differences in the millionth's decimal place when computing probability ratios.

1. *AMT re-enactment*. The mastery cut-off was converted to Θ_c using the test characteristic curve (see [14], formula {24}) and the item parameter database constructed from the respective parameter estimation sample (either $n = 25$ or 50). The value of Θ_c was used as the initial prior Θ and the prior variance was set to one, as recommended by Weiss and Kingsbury [9]. The MISS procedure was used to select the next test item for the re-enactment for each examinee (see [14], formulas {10} to {12}). The correctness of the examinee's response to that item was determined by retrieving it from the database. Bayesian updating of Θ and its variance was accomplished with Owen's method [17] (see [14], formulas {13} to {22}). After each item was "administered", the AMT stopping rules were applied using a .95 confidence interval (see [14], formulas {25.1} to {25.3}). If a decision could be reached, the re-enactment was ended at that point. The number of questions answered correctly and incorrectly in the AMT and the decision reached for that examinee were written to a computer data file. Also stored in that file were the total test score for that examinee and the agreement between the AMT decision and the total test decision. If no decision could be reached by the AMT model before exhausting the test item pool, then a decision was forced at the end of the test: If the current estimate of Θ was greater than or equal to Θ_c, the examinee was considered to be a master; otherwise a nonmaster.

2. *SPRT*. The mastery and nonmastery levels required by the SPRT were empirically established from the parameter estimation samples, as described above. Since the SPRT requires random selection of items, test items were "administered" in a random order. Alpha and β levels were set at 0.025, to make the overall decision error rate (.05) equivalent to the .95 confidence interval method used in the AMT approach. When the SPRT reached a mastery or nonmastery decision, results were stored in a separate data file in the same manner as described above for the AMT.

3. *EXSPRT-R*. As in the SPRT, items were "administered" in a random order. However, the rule bases constructed from the parameter estimation samples were used, of course, in the EXSPRT-R method of Bayesian updating (formula {2}, with equal prior probabilities) and SPRT stopping rules ({3.1} to {3.3}). For a description of EXSPRT-R procedures, see [12] for an example of expert systems reasoning during computer-based testing. When the EXSPRT-R reached a decision, the test re-enactment was ended and results written to a data file as before.

4. *EXSPRT-I*. This method was the same as the EXSPRT-R, except that items were selected intelligently, based on their utility indices (see formulas {4} to {8}). Thus, like the AMT, items were not "administered" randomly for each re-enactment. Since no feedback was given during the test it is unlikely that decisions reached by both AMT and EXSPRT-I methods would be systematically affected by factors other than differences in the adaptive methods themselves. One mitigating factor might be examinee fatigue, where they were more likely to answer questions incorrectly at the end of the long and difficult test. However, since all test items were originally administered in a different random order for each individual, it is very unlikely that fatigue would systematically bias any findings.

Results from the First Study

For the DAL test, IRT item parameters (b_i's) were estimated from samples of 25 and 50 examinees. EXSPRT rule bases were also derived from the same samples. Descriptive information is given about the two samples in the left side of Table 1. It can be seen that there were about the same proportions of masters and nonmasters in each sample. In the sample of 50 there were 23 masters whose average test score was 87.3 percent, and 27 nonmasters who scored 45.1 percent correct.

Mean test lengths of each of the four methods, variation in test lengths, and decision accuracies were compared. If the decision made by an adaptive method was the same as that reached on the basis of the entire test item pool, this was considered to be a "hit". Thus, the accuracy measures are the percent of correct predictions made by each method. There were 28 nonmasters and 25 masters identified by the entire 97-item test, when the cut-off score was set at 72.5 percent correct.

First, note that the parameter sample size seems to make little difference in the mean test length *within* each method. For example, within the AMT model 20.6 items were required for nonmastery decisions when item parameters were based on a sample of 25, compared to a mean of 18.3 for the sample of 50. For the EXSPRT-I, 5.6 items were required for nonmastery decisions in the sample of 25, compared to a mean of 5.9 for the parameter sample of 50. Please note--and this is confusing--that the mean test lengths for each of the four methods are based on the same 53 test administrations, where all 97 items were originally given, and which were re-enacted under each adaptive method. The size of the *parameter estimation* sample refers to the number of examinees randomly selected on whom the item difficulties were estimated for the AMT model and on whom the item rule bases were constructed for the EXSPRT-R and EXSPRT-I models.

Decision accuracies. For the 53 administrations of this DAL test there does seem to be some difference in decision accuracies within each model for the two parameter estimation sample sizes. The decision accuracies tended to be high for all methods. Decision accuracies were compared to expected values of .975 correct mastery decisions and .975 correct nonmastery decisions, using Chi-square goodness of fit tests (e.g., see [18]). A significant Chi-square ($p < .05$) means that the observed decision accuracies *departed* from what was expected according the *a priori* decision error rates that were established for each of the four adaptive testing methods.

When 25 examinees were used for parameter estimation, there were two significant departures from expected accuracy. The EXSPRT-I was 85.7 percent accurate in nonmastery decisions, which significantly differed from the expected 97.5

Table 1. Efficiency and Accuracy of the Four Adaptive Testing Methods in the First Study.[6]

Item Parameter Sample Description		ADAPTIVE TESTING METHOD			
		AMT	SPRT	EXSPRT-R	EXSPRT-I
	Mean Score (S.D.) n	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy
Masters	87.46 (7.90) 12	8.40 (9.62) 100.0	8.72 (5.16) 92.0	7.56 (3.22) 100.0	5.44 (1.23) 100.0
Nonmasters	42.66 (15.83) 13	20.57 (24.45) 96.4	10.54 (7.14) 85.7*	12.71 (15.46) 96.4	5.64 (2.02) 85.7*
Total	64.16 (25.99) 25	14.83 (19.77) 98.1	9.68 (6.29) 88.7	10.28 (11.65) 98.1	5.55 (1.68) 92.5
Masters	87.27 (7.89) 23	8.28 (8.19) 100.0	10.36 (6.92) 96.0	8.44 (5.74) 96.0	6.84 (2.64) 100.0
Nonmasters	45.06 (16.25) 27	18.29 (24.43) 96.4	10.11 (10.97) 89.3*	9.39 (9.15) 92.9	5.93 (2.28) 92.9
Total	64.47 (24.89) 50	13.57 (19.14) 98.1	10.23 (9.20) 92.5	8.94 (8.94) 94.3	6.36 (2.47) 96.2

*Percent accuracies were tested by goodness of fit, where .975 accuracy was expected according to the *a priori* error rates for masters and nonmasters. Only those percent accuracies which differed significantly from the expected accuracies, according to a chi-square test (d.f. = 1, *p* < .05) are marked with an asterisk.

---

[6] Alpha = ß = 0.025 for the SPRT, EXSPRT-R, and EXSPRT-I; a .95 confidence interval was used with the AMT. There were 53 administrations of the DAL test which were re-enacted for each of the four adaptive methods.

percent accuracy. At the same time, however, the EXSPRT-I was reaching decisions when the other models were requiring two to four times as many items. The SPRT accuracy for nonmastery decisions was also significantly lower than expected.

When 50 examinees were used for parameter estimation, the AMT, EXSPRT-R, and EXSPRT-I models were within the expected range of accuracy. The SPRT failed to make as many correct nonmastery decisions as were expected. What is notable is how well *all* of the adaptive methods predicted total test decisions, while using between 5 and 20 items from the 97-item pool to reach those decisions--a very substantial reduction in test lengths (95 to 80 percent decrease).

Efficiency. A repeated measures ANOVA (MANOVA) was conducted to see if there were significant differences among the mean test lengths for the four adaptive methods. This was done for the results based on the parameter sample of 50 for the 53 test administrations. Hotelling's T^2 was significant at the .05 level. However, the sphericity assumption was violated, due to the large differences in variances among the four methods. A *post hoc* comparison procedure suggested by Marascuilo and Levin ([19], pp. 373-381) for this kind of situation was conducted for all pair-wise contrasts of mean test lengths. One statistically significant difference was found. The mean test length for the SPRT was significantly greater than that for the EXSPRT-I. Even though some of the other contrasts have greater magnitudes of difference, the within-method variances are very different themselves. It can be noted that, overall, the AMT model required about twice as many items to reach decisions (13.6) as did the EXSPRT-I (6.4), though it was not statistically significant at the .05 level.

The variances in average test lengths within each adaptive method were significantly different, as noted above in violation of the sphericity assumption. The variance in test lengths for the AMT model was approximately 60 times larger than that for the EXSPRT-I model (19.14^2 vs. 2.47^2). In the AMT model, tests tended to be longer before nonmastery decisions were reached, and there was much more variation in test lengths compared to the remaining models. The variation in lengths of tests with EXSPRT-I method was relatively small compared to variation in the remaining models.

SECOND STUDY

Computer Functions Test

A computer-based test on how computers work, consisting of 85 items, was constructed. The COM test, as it is referred to here, was comprised of about half multiple-choice, one-fourth binary choice, and one-fourth fill-in type questions (Cronbach α = .94). Compared to the DAL test, the COM test was much easier for most examinees (mean score = 79.0 percent, S.D. = 13.6).

Examinees

About half of those who took the COM test were from two sections of an introductory graduate-level course on use of computers in education. The remainder were mostly volunteers from an undergraduate-level course for non-education majors who were learning to use computers. A small number of students were volunteers recruited at the main library on campus.

Test Administration and Experimental Methods

The COM test was individually administered by the Indiana Testing System in the same manner as the DAL test. There were a total of 104 administrations of the COM test in the second study. The same four adaptive testing methods were reenacted from actual examinee test data in the very same manner as described above for the DAL test.

Results from the Second Study

Since there were more administrations of the COM test, parameter estimation samples of 25, 50, 75 and 100 were selected at random. Four sets of b_i coefficients were obtained for the AMT model and four rule bases were constructed for the EXSPRT models based on the same four parameter estimation samples. See the left sides of Tables 2.1 and 2.2 for descriptive information about the parameter estimation samples.

Accuracy of predictions. When the parameter estimation sample was 25, all four adaptive methods did not perform as well as expected in correctly predicting nonmasters in the 104 administrations of the COM test. Chi-square goodness of fit tests showed that all four methods significantly departed from the expected accuracy rates. EXSPRT-I had the worst accuracy, but it should be noted that there were only *seven* nonmasters in the estimation sample for creating the rule base, so this is not surprising.

When the parameter estimation sample was 50, the AMT and EXSPRT-I models still made significantly fewer correct nonmastery decisions than expected *a priori*. On the other hand, the SPRT and EXSPRT--both of which use random selection of items vs. intelligent selection in the AMT and EXSPRT-I--predicted masters and nonmasters correctly within the bounds of expected error rates.

When the parameter estimation sample was 75 (55 masters and 20 nonmasters when the cut-off was 72.5 percent correct), all models predicted well except the AMT, which made significantly fewer correct nonmastery decisions than were expected *a priori*.

When the parameter estimation sample was 100, the AMT model still had problems with accuracy of nonmastery classifications. And strangely enough, the AMT model also made significantly fewer correct *mastery* decisions than were expected. The SPRT, EXSPRT-R, and EXSPRT-I all correctly predicted masters and nonmasters within the bounds of expected accuracies. It should be noted that it is generally recommended that a minimum of 200 examinees be used for estimating b_i parameters in the IRT-based, one-parameter AMT model. Only half that number were available in this study. Thus, it is not surprising that the AMT model performed less well than it should, since estimation of the item difficulty parameters was not as precise as desired.

Efficiency. Average test lengths of the four adaptive methods were compared for the 100 examinee parameter estimation situation only. See the bottom half of Table 2.2. A MANOVA again revealed that the sphericity assumption was violated, and so the same procedure as described above for the DAL test was used in post hoc comparisons of the adaptive COM test length means [19].

Table 2.1. Efficiency and Accuracy of the Four Adaptive Testing Methods in the Second Study.[7]

Item Parameter Sample Description		ADAPTIVE TESTING METHOD			
		AMT	SPRT	EXSPRT-R	EXSPRT-I
	Mean Score (S.D.) n	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy
Masters	86.21 (6.35) 18	8.37 (13.59) 97.4	11.71 (7.80) 98.7	10.05 (5.42) 98.7	4.57 (3.60) 98.7
Nonmasters	48.07 (9.10) 7	33.93 (31.83) 82.1*	14.39 (15.81) 85.7*	15.00 (10.41) 82.1*	7.07 (2.36) 67.9*
Total	75.53 (18.84) 25	15.25 (23.02) 93.3	12.43 (10.55) 95.2	11.38 (7.39) 94.2	5.24 (3.49) 90.4
Masters	87.16 (5.68) 35	11.83 (18.23) 94.7	15.08 (9.06) 96.1	11.71 (8.28) 98.7	5.72 (3.92) 96.1
Nonmasters	53.65 (10.44) 15	31.89 (29.97) 78.6*	17.39 (14.50) 92.9	15.82 (12.70) 96.4	7.93 (6.21) 89.3*
Total	77.11 (17.15) 50	17.23 (23.61) 90.4	15.70 (10.77) 95.2	12.82 (9.78) 98.1	6.32 (4.72) 94.2

*Percent accuracies were tested by goodness of fit, where .975 accuracy was expected according to the *a priori* error rates for masters and nonmasters. Only those percent accuracies which differed significantly from the expected accuracies, according to a chi-square test (d.f. = 1, *p* < .05) are marked with an asterisk.

BEST COPY AVAILABLE

---

[7] Alpha = β = 0.025 for the SPRT, EXSPRT-R, and EXSPRT-I; a .95 confidence interval was used with the AMT. There were 104 administrations of the COM test which were re-enacted for each of the four adaptive methods.

Table 2.2. Efficiency and Accuracy of the Four Adaptive Testing Methods in the Second Study (cont'd).

Item Parameter Sample Description		ADAPTIVE TESTING METHOD			
		AMT	SPRT	EXSPRT-R	EXSPRT-I
	Mean Score (S.D.) n	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy	Mean Length (S.D.) Accuracy
Masters	87.68 (5.93) 55	10.21 (16.96) 94.7	16.64 (10.35) 97.4	11.78 (6.34) 97.4	7.70 (7.13) 94.7
Nonmasters	56.00 (10.17) 20	28.93 (28.58) 82.1*	16.29 (16.96) 92.9	14.75 (15.54) 100.0	7.82 (4.85) 100.0
Total	79.23 (15.84) 75	15.25 (22.21) 91.3	16.55 (12.39) 96.2	12.58 (9.71) 98.1	7.73 (6.57) 96.2
Masters	87.47 (6.33) 75	13.75 (20.80) 93.4*	16.97 (10.75) 96.1	13.58 (9.51) 98.7	7.64 (6.12) 94.7
Nonmasters	56.00 (11.34) 25	31.50 (29.77) 78.6*	13.04 (10.66) 92.9	12.32 (10.78) 96.4	8.93 (7.41) 100.0
Total	79.60 (15.77) 100	18.53 (24.70) 89.4	15.91 (10.82) 95.2	13.24 (9.83) 98.1	7.99 (6.48) 96.2

*Percent accuracies were tested by goodness of fit, where .975 accuracy was expected according to the *a priori* error rates for masters and nonmasters. Only those percent accuracies which differed significantly from the expected accuracies, according to a chi-square test (d.f. = 1, *p* < .05) are marked with an asterisk.

BEST COPY AVAILABLE

When nonmastery decisions were made, the AMT model required significantly longer tests than either the SPRT, EXSPRT-R or EXSPRT-I. The AMT model required about 32 items to reach nonmastery decisions, compared to the EXSPRT-I, which required about nine items. Moreover, the AMT made significantly fewer correct nonmastery decisions than expected, as noted above. When mastery decisions were reached, test lengths for the SPRT and EXSPRT-R methods (15 and 12) were significantly longer than the EXSPRT-I (6 items). Mean test lengths for mastery decisions in the AMT and EXSPRT-I models were not significantly different at the .05 level.

When looking at decisions overall, the following contrasts were significantly different: the AMT, SPRT, and EXSPRT-R methods each required significantly longer tests than did the EXSPRT-I model. The AMT model required over twice as many items as did the EXSPRT-I (19 vs. 8).

Summary. It would appear from the COM test data that the EXSPRT-I is significantly more efficient than the other adaptive methods. Indeed, it is rather remarkable that the EXSPRT-I can make such highly accurate mastery and nonmastery decisions with relatively few test questions. It is also notable that the EXSPRT-R and SPRT also made highly accurate predictions, but were less efficient than the EXSPRT-I. The AMT performed worst of all, not only resulting in longer adaptive tests but also in making significantly more prediction errors than theoretically expected.

DISCUSSION

Adaptive tests tended to be shorter with the DAL test than with the COM test. See Tables 1, 2.1 and 2.2. Of the 53 administrations of the DAL test, there were 28 nonmasters and 25 masters when the cut-off was set at 72.5 percent and when examinees answered all 97 items. The overall average test score was 63.2 (S.D. = 24.6). In the second study with the COM test there were 104 administrations of this test, with 76 masters and 28 nonmasters when the entire 85-item test was taken (grand mean = 79.0, S.D. = 13.6, mastery cut-off = 72.5 percent).

A similar study was conducted by Plew [13] with yet a different test on computer literacy (referred to as the LIT test here). In his sample of 183 examinees there were 54 masters and 129 nonmasters based on total test results from the 55-item pool. The cut-off for this test was 59.5 percent, and the overall average score was 51.5 percent (S.D. = 14.2).

One thing that appears to affect the average test lengths is the location and shape of the distribution of examinee achievement levels in relation to the cut-off selected. In the first study, the distribution was somewhat bimodal and relatively flat, with about half the examinees scoring above and below the cut-off. In the second study, the distribution was positively skewed, with about three-fourths of the examinees scoring above the 72.5 percent cut-off on the entire test item pool. In the Plew study, over two-thirds of the examinees were classified as nonmasters on the entire 55-item test [13]. The distribution of this group was close to normal, with the mean being about 8 percentage points below the selected cut-off.

The author has previously conducted a number of computer simulations comparing the three-parameter AMT model with the SPRT and a third adaptive method based on Bayesian posterior beta distributions ([14],[20]). One important finding in those studies was that none of the adaptive methods performed as well as expected--and average test lengths tended to be longer--when the distribution of examinees was mostly clustered around the cut-off. Adaptive tests were shorter and accuracies agreed with theoretical expectations when the distributions of examinee achievement levels were much flatter. The same phenomenon appears to have occurred in the present two empirical studies, as well as in Plew's.

The second factor that may affect results is the number of test items in each pool and their properties. When there are more test items, and there are more items available at each ability or achievement level, then both the AMT and EXSPRT-I tend to be more efficient and more accurate. In both adaptive methods which rely on "intelligent" selection of items, Bayesian posterior estimates are affected more dramatically when there are highly discriminating items available whose difficulty levels are close to the current estimate of an examinee's achievement level. A real problem occurs with smaller item pools, as was the case with the LIT test in Plew's study: After the best items have been administered early in a test, the remaining items tend to provide little additional information. That is, there are diminishing returns after some point because there are no really appropriate items left.

SUMMARY

Expert systems can be used to aid decision makers. A computerized adaptive test (CAT) is one kind of expert system, though not commonly recognized as such. When item response theory is used in a CAT, then the knowledge or rule base is a set of item characteristic curves (ICC's).

Normally an expert system consists of a set of questions and a rule base. An inference engine uses answers to the questions and the rule base to choose from a set of discrete alternatives. If an adaptive test is viewed this way, then it is possible to construct "If ..., then ..." rules about test items that are not functions, as are ICC's. A new approach, termed EXSPRT, was devised that combines expert systems reasoning and sequential probability ratio test stopping rules. EXSPRT-R uses random selection of test items, whereas EXSPRT-I incorporates an intelligent selection procedure based on item utility coefficients.

These two new methods were compared to the traditional SPRT and to an IRT-based approach to adaptive mastery testing (AMT). Two empirical studies with different tests and types of examinees were carried out.

In the first study the EXSPRT-I model required about half as many items as did the AMT approach (6 vs. 14), though the difference was not statistically significant. When 50 examinees were used for item parameter estimation and rule base construction, all four methods (AMT, SPRT, EXSPRT-R and EXSPRT-I) made highly accurate mastery and nonmastery decisions.

In the second study the EXSPRT-I method again required about half as many items as did the AMT model (8 vs. 19), and this time the difference was statistically significant. When 100 examinees were used for estimation purposes, the SPRT, EXSPRT-R, and EXSPRT-I correctly predicted masters and nonmasters within the bounds of the expected theoretical error rates. The AMT model, however, made significantly more prediction errors than expected.

Although further research is needed, the EXSPRT-I initially appears to be a strong alternative to both IRT- and SPRT-based adaptive testing when categorical decisions about examinees are desired. The EXSPRT-I is clearly less complex than IRT, both conceptually and mathematically. It also appears to require many fewer examinees to establish empirically a rule base when compared to the large numbers required to estimate parameters for item response functions in the IRT model.

On the other hand, the EXSPRT is vulnerable, as is classical test theory, in that a representative sample of examinees must be selected for constructing rule quadruplets. This seems to be a small price to pay for the advantages of theoretical parsimony and operational efficiency.

REFERENCES

1. Winston, P. and Prendergast, K. (1984). *The AI business: The commercial uses of artificial intelligence*. Cambridge, MA: The MIT Press, 17-40.
2. Davis, R. (1984). Amplifying expertise with expert systems. In P. Winston and K. Prendergast (Eds.), *The AI business: The commercial uses of artificial intelligence*. Cambridge, MA: The MIT Press, 17-40.
3. Kearsley, G. (Ed.), (1987). *Artificial intelligence and instruction: Applications and methods*. Reading, MA: Addison-Wesley.
4. Sleeman D. and Brown, J. (Eds.) (1982). *Intelligent tutoring systems*. New York, NY: Academic Press.
5. Clancey, W. (1987). Methodology for building an intelligent tutoring system. In G. Kearsley (Ed.), *Artificial intelligence and instruction: Applications and methods*. Reading, MA: Addison-Wesley, 193-227.
6. Weiss, D. and Kingsbury, G. (1984). Application of computerized adaptive testing to education problems. *Journal of Educational Measurement, 21*, 361-375.
7. Heines, J. (1983). Basic concepts in knowledge-based systems. *Machine-Mediated Learning, 1*(1), 65-95.
8. Wright, B. (1977). Solving measurement problems with the Rasch model. *Journal of Educational Measurement*, 14(2), 97-116.
9. Frick, T. (1989). Bayesian adaptation in computer-based tests and computer-guided practice exercises. *Journal of Educational Computing Research*, 5(*1*), 89-114.
10. Schmitt, S. (1969). *Measuring uncertainty*. Reading, MA: Addison-Wesley.
11. Wald, A. (1947). *Sequential analysis*. New York: Wiley.
12. Frick, T. (1990). Analysis of patterns in time: A method of recording and quantifying temporal relations in education. *American Educational Research Journal*, 27(1), 180-204.
13. Plew, G. T. (1989). *A comparison of major adaptive testing strategies and an expert systems approach*. Bloomington, IN: Doctoral dissertation, Indiana University Graduate School.
14. Frick, T. (1990). A comparison of three decision models for adapting the length of computer-based mastery tests. *Journal of Educational Computing Research*, *6*(4), 479-513.

15. Frick, T. (1986). The Indiana Testing System (ITS, Version 1.0). Bloomington: Department of Instructional Systems Technology, School of Education, Indiana University.
16. Mead, R., Wright, B., and Bell, S. (1979). BICAL (Version 3). Chicago: Department of Education, University of Chicago.
17. Owen, R. J. (1975). A Bayesian sequential procedure for quantal response in the context of adaptive mental testing. *Journal of the American Statistical Association*, *70*, 351-356.
18. Glass, G. and Hopkins, K. (1984). *Statistical methods in education and psychology*. Englewood Cliffs, NJ: Prentice-Hall.
19. Marascuilo L. and Levin, J. (1983). *Multivariate statistics in the social sciences: A researcher's guide*. Monterey, CA: Brooks/Cole.
20. Frick, T., Luk, H.-K., and Tyan, N.-C. (1987). *A comparison of three adaptive decision-making methodologies used in computer-based instruction and testing*. Bloomington, IN: Final Report, Proffitt Foundation, Indiana University School of Education.

Title:

Qualitative Research Methodology Issues

Author:

Robert G. George

QUALITATIVE RESEARCH METHODOLOGY ISSUES

Orlando, Florida
February 16, 1991
Robert G. George

Current movements suggest that qualitative or ethnographic research will be the major research trend going into the 21st Century. Many researchers, however, are finding it difficult to support these methods of research because of the strong emphasis placed on quatitative research which still influences their thinking.

In order to facilitate communication there needs to be a definition of qualitative research. There appears to be as many qualitative methodologies and definitions as there are authors. A sample of terms follows:

Naturalistic
Phenomenological
Interpretive
Ethnographic
Field Study
Case Study
Ecological Psychology
Holistic Ethnography
Cognitive Anthropology
Symbolic Interaction
Ethnography of Communication
Participant Observation

The above terms do not exhaust the variety of methodologies one encounters in reading. With such a great variety of terms, one general definition for purposes of discussion within the confines of this paper necessarily has to be very broad and general in nature.

Qualitative research requires that the researchers accept certain philosophical assumptions in accordance with the methodology being used. Moreover, these philosophical assumptions need not be totally objective nor quantitative in nature. The data collected for such research is always verabally oriented, but may also contain numerical data. In fact, it is the verbally oriented data that most clearly separates the realms of quantitative and qualitative.

Qualitative Research Issues

New or Fad: It could be argued that researchers were qualitative in nature long before they were quantitative. In fact the 1930's were the years that ushered in the quantitative people. In fact, early experimental people had trouble getting published. But the quantitative movement was so strong as to push qualitative studies into disfavor at first, and later into disdain by the educational community. Now there has been a movement back to qualitative and the younger researchers are anxiouly joining the throngs of the new wave which is also pulling older researchers away from their prior efforts in order to not be deemed as old fashioned or out of date. In the rush for "newness" creative researchers are coining their own phrases, methodologies, standards, and philosphies. When you think of the years of study by someone like Jane Goodall on her subjects, you can better appreciate the humor in the article by Ray Rist (1980) entitled "Blitzkrieg Ethnography: On the Tranformation of a Method into a Movement." Goodall took years and blitzkrieg took days. Perhaps there is a rush to create your own qualitative methodology in hopes of attaining a certain immortality.

Too Many Methodologies: If new fields of endeavor become breeding grounds for creative methods, then it should not surprise anyone when the number of methodologies reach a point where research readers become frustrated over the inability to fully grasp the understanding of every new methodology. This does not mean that all research readers understand all quantitative methods, but most of the quantitative methods are readily available for study in textbooks and are rather formal in nature. Many of the new methods of qualitative research are created during or after the completion of the study. Then the next researcher who wishes to duplicate or to incorporate the new methodolgy into his study will be fortunate if he can capture all of the nuances of the new method. It is almost like every qualitative researcher must devote some space to alerting the reader to the researchers' perception of the methodology used in the study. Perhaps time will start to create standards.

Lack of Standards for Data: It would stand to reason that if the basic assumptions, the terms, and the methodologies are not standardized, then you can appreciate that the analysis of data has not been standardized. This does not mean that attempts are not being made to standardize data treatment. In fact, whole books such as the one by Miles and Huberman (1984) have been written dealing with qualitative data analysis. Several textbooks on research such as McMillan and Schumacher (1989) and Hopkins and Antes (1990) have been written that discuss data collection as a part of the section on qualitative research. But the research community has not established hard and fast rules on analysis. Moreover, the major journals of research have not established any firm guidelines for publication. Some people are concerned that the ability to write persuasively might lead to the acceptance of form over substance.

Questionable Samples: Quantitative researchers have been heard to jokingly say that all of research is a question of sampling. Sampling the literature, sampling the subjects, sampling the data, and perhaps sampling the data to be reported. But lying beneath the surface of the joke is the thought that the researcher is trying to get a "true" representation of the population under study. Accepted methods of sampling, albeit often knowingly violated, are considered to render the best representation of the population. Now comes qualitative research methods that do not place sampling restraints upon the researcher. This is especially true of the sampling of subjects. Hence subjects may or may not be representative of a larger population. The qualitative researcher at this point may be attempting to remind this author that many qualitative researchers do not feel a need to project his findings beyond the immediate subjects under study. That each population, no matter the size, might be the whole of the interest of the researcher.

Contamination of Subjects: Qualitative researchers necessarily are present as an outside or foreign presence. There are various levels of intervention, depending upon the method of qualitative research being done, but whatever the level of presence the question of affecting behavior has to be asked. Would the events of an observed situation unfold in the same manner regardless of the presence of a researcher? I suspect the ability of the researcher would be the major determinent of the level of contamination.

Inability to Project Findings: This brings the reader to a real problem that is basic to all research, the adding of new knowledge to our base of knowledge. Qualitative research can tell you what the current state of affairs is at a particular point in time, but there are limiting conditions in generalizing to other supposedly similar situations. If the researcher and the reader of the research are only interested in a situation for the sake of understanding, then qualitative research is apropos. If, however, that researcher is trying to change behavior or predict behavior he cannot use qualitative research findings. Qualitative research cannot be used to predict behavior under current educational research standards.

A New Philosophy Needed: Finally, the key issue that evades most discussions regarding qualitative and quantitative research is philosophy. There is a definite need to develop a strong philosophy by the researcher before he enters into a qualitative study. The various qualitative methods require various philosophical backgrounds before embarking upon their use, and the philosophical difference are great enough to cause great differences in the acceptance of findings.

Obviously, the differences in the philosophies of quantitative and qualitative research methods are great enough to cause open rejection of findings one from the other. The findings of Piaget were based upon qualitative research, but none of his writings were accepted until his theories had been tested by quantitative studies in the United States. The question arises: "Were Piaget's theories any less valid when based upon qualitative research methods than when they were posited based upon quantitative methods?" The seeking of truth is not undesireable in any philosophy, but the definition of truth may vary. Using qualitative research will require new ways of viewing knowledge and people.

What will happen as the philosophies of qualitative and quantitative research begin to clash? There is reason to believe that the core of each camp will remain unchanged in their philosophies. That there will be some researchers daring enough to venture back and forth between the two camps is to be expected. But there will also be those charlatans who will use and abuse qualitative research methodologies in hopes of circumventing research standards that appear too rigid or too mathmatical in nature. As O'Conner (1990) pointed out in his paper, "...too often researchers cover over their failure to sort out what they are trying to do by hiding behind the early ambiguity of qualitative research projects--using the old line "I won't know unti I'm in the field." It appears that it will be the task of the research community to be willing to take the time to consciously read and evaluate qualitative reports with the willingness to separate the charlatans from the serious endeavors.

Summary

What was discussed here today was a list of issues that the research community should address in some manner. Newness or fads are always concepts that attract, but new concepts and fads should be refined and improved upon with use. Several issues presented in this paper should be studied further to see how the research community will deal with them. Certainly standards for data, sample problems, subject contamination, and the use of the findings from qualitative research should be studied for possible refinement.

Perhaps the most important issue, a new philosophy of knowledge, is the one that will be the last to be dealt with by researchers. Researchers are terribly guilty of asking questions just because they think the questions have not been asked before. Perhaps the philosophical approach would ask why the question was asked. Perhaps the questions themselves will need to be evaluated in qualitative terms.

This paper did not offer nor intend to be the definitive statement about qualitative research. This paper was intended to evoke discussion in the research community about the research trend of the twenty-first century.

References

Hopkins, Charles D. and Antes, Richard L. (1990). Educational Research: A Structure for Inquiry. Third Edition. Itasca, Illinois: Peacock Press.

McMillan, James H., and Schumacher, Sally (1989). Research in Education: A Conceptual Introduction. Second Edition. Glenview, Illinois: Scott, Forsman and Company.

Miles, Matthew B., and Huberman, A. Michael (1984). Qualitative Data Analysis: A Sourcebook of New Methods. Beverly Hills, California: Sage Publications.

O'Connor, Terrence. (1990) Knowing What We Really Know. Paper presented at a Qualitative Forum, Indiana State University, Terre Haute, Indiana. Unpublished.

Rist, Ray C. (1980) Blitzkrieg Ethnography: On the Transformation of a Method into a Movement. Educational Researcher, February, 8-10.

Title:

Supporting Learning with Process Tools: Theory and Design Issues

Authors:

David A. Goodrum
Randy A. Knuth

Introduction

We have been working to develop our understanding, in terms of instructional design and development, of the implications of a constructivist epistemology (Brown,1989). From this perspective, learning is seen as a *constructive* rather than *accumulating* process. That is, people actively construct knowledge and understanding through their individual and social experience in, and interaction with, the world (Streibel, 1986). According to Resnick (1985, p. 2570), "These constructions respond to information and stimuli in the environment, but they do not copy or mirror them." Every concept that we learn derives its meaning through its relationship to the context in which it is intimately embedded and from the tasks in which it is used. Brown and his colleagues have referred to this notion as situated cognition (Brown, Collins, & Duguid, 1989).

We believe that the design of learning environments that afford opportunities to experience concepts embedded in authentic contexts and authentic tasks must embody two characteristics: 1) a rich database representing both core content and its related context; and 2) support for cognitive processes, both individual and social. These components have historically been supported through a strategy of human contact and more recently through computer technologies.

The Tension Between Content and Process

We have seen, however, a lopsided investment of technological energy and instructional focus into one of these components: the content database. The major area of effort has been in the delivery of a specific core and contextual database along with practice strategies aimed at student retention and understanding of the database. Computers and other teaching machines have seemed to offer an efficient and consistent method to deliver sequenced material with a navigation strategy of branching and co-requisite learning activities such as tutorials, simulation, and drill and practice. Most recently, hypertext and hypermedia seem to offer the designer's dream in putting at the learner's disposal unheard of amounts of potential source material along with a navigational strategy of browsing and, when that fails, guided tours.

The focus on the content database has occurred, we believe, because of a belief about the nature of learning which is often contrasted with constructivism. Instructional designers have traditionally focused on creating learning environments intended to transmit a set of information to the learner, for example through appropriate displays and sequences (Merrill, 1983) which promote *individual* learning. New technologies, for example hypermedia environments, have, by and large, been based on this same view of learning. Hypermedia is thought to provide learners with the ability to *browse* through information spaces, acquiring information and knowledge as a result of their journey (Byers, 1987). The learner's goal is to *find the particular information* that he or she is to know. The design of hypermedia systems has centered on the creation of large navigable databases.

Unless special tools are made available to help the learner actively construct knowledge, hypermedia databases might actually encourage the learner to be a passive receiver of information. Passive browsing of an information space (i.e., text) results in superficial learning (Brown, 1981). Additionally, typical hypermedia systems provide little or no feedback (Hammond, 1989) — a critical component to the traditional approach to instruction. In contrast, good readers are active readers who utilize a variety of strategies or processes for managing their understanding of the material (Garner,1987). For example, they take notes, highlight important passages, make outlines, mark text to be reread, paraphrase summaries, and so on. Processes shown to promote even higher levels of understanding include designing a presentation, writing a synthesis paper, keeping a journal, and using a critical analysis strategy to analyze the material. An important relationship exists between the artifacts created from doing authentic tasks in authentic contexts and the learner's constructed knowledge.

The development of learning environments based on the individual learning and information transmitting paradigm also does little justice to take advantage of the social nature of learning. We have come to value the notion that not only is learning a constructive process but also that meaning is negotiated through collaboration with others. Each of the processes listed in the preceding paragraphs is enhanced through collaboration. In Vygotsky's view (1978), an individual is limited in the problems that can be solved and the tasks that can be completed on one's own. Help from others can provide a way for going beyond one's current level of competence and working on problems in one's *zone of proximal development* (ZPD). An expert, for example, can include an individual learner's attempts in accomplishing a task and provide feedback. In this way, the individual can perceive the appropriateness of the attempts and can come to see the task as the more expert person sees it. It is this social interaction with the expertise of others that assists the individual in actively constructing an understanding of the problem domain (Newman, 1990).

It is our assumption, therefore, that the path from novice to expert is accomplished by an active engagement with the important concepts *in concert with others*. As is commonly believed, a good way to come to understand a concept fully is to teach it to someone else. Part of this is the preparation a teacher or tutor goes through with content materials. Teachers report having learned more about both the content and the abilities of their learners through social interaction with their students (Cobb, & Steffe, 1983). Clearly, learning for both novice and expert, requires substantially more than simply 'seeing' the concepts. Unfortunately, the common technique of having individuals browse through a hypermedia database encourages the learner to stay at the novice level.

Process Tools

What constructivist ideas make evident is that we have paid too little attention to the cognitive and social processes that are an equally inherent and important part of learning. We have, therefore, shifted our focus to the identification of learning processes and to building tools – that is, *process tools* – designed to support these processes. We don't dispute the value of employing database technology in the educational setting. What we object to is the lack of attention given to the development of process tools and their integration with content databases.

We anticipate that process tools will also have an advantage over traditional approaches to instruction in the area of motivation: people inherently value that which they themselves construct. Learners commonly evaluate didactic instruction based on what they received, or less tactfully, on what it did *to* them. Our contention is that this reaction is encouraged by the passive nature of the instructional presentation. Similarly, in many extant educational applications of hypermedia systems, people are encouraged to browse through information, sometimes with explicit goals, sometimes without them (Hammond, 1989). They then leave the experience with no product — perhaps a few memories, perchance some incidental learning — and often just a score on an exam. Process tools, on the other hand, are designed for use by learners to collaboratively construct a product of which they have intimate ownership.

Computer Supported Work

We have found interesting research and development in the area of process tools, although they are generally not referred to as such. However, this research and development is occurring in disciplines from which either educators or instructional developer's seldom draw. Designers in the areas of group decision support systems (GDSS) (e.g., DeSanctis & Gallupe, 1987) and computer supported cooperative work (CSCW) (e.g., Greif, 1988) have developed exciting electronic tools to support group processes, usually for business , research, or military operations. In fact, we have borrowed ideas from this work for our project. Borrowed ideas, however, are useful only in their ability to transfer from one domain to another. Some portions of the analogy will inevitably fall short. We have found that GDSS and CSCW analogies map very well to our domain at the level of systems design and even group interaction technology. Where it does not seem to apply in an educational domain is in its focus and purpose. The systems developed by these fields are primarily intended to help people in business and research situations accomplish tasks. In the fields of education and instructional technology the primary focus, in light of a constructivist epistemology, is to help people *learn* how to accomplish tasks.

A subtle but important difference exists between the intended purposes of tools in work and learning settings. It is necessary for both workers and students to be involved in authentic tasks. Process tools need to make more efficient and explicit the learning and social processes so that students do not have to flounder in trial and error. Process tools for business, on the other hand, need to help workers get the job done more efficiently regardless of how much learning occurs. This is the dilemma that instructional designer's faces when trying to create or find authentic tasks: Taking the tenets of situatedness to their logical conclusion would be to simply send students to work. But, since the primary purpose of work is to get the job done, businesses attempt first to employ people who *already know how* – who, in a sense, are*experts* – to accomplish the task. Learning that naturally occurs as a result of doing the task becomes secondary to getting the task done. Thus, sending students to work does not work for the business and does not work for the student; schooling needs to make the learning processes explicit, that is encourage learners to be reflexive about their learning (Cunningham, 1987), whereas business can afford less tolerance for learning at the expense of getting the job done.

In other words, the crux about what "business" is about is not learning but in accomplishing (doing) tasks as the *primary* activity. Furthermore, groups are brought together because *physical proximity* is the strategy (i.e., *technology*) that gets things done (Kraut, 1989). In contrast, education is about *learning to do* a task, about making people better thinkers. In education, *learning* is the primary activity. Constructivist theory explicitly suggests that the *strategy*, or, again the *technology*, that best accomplishes this is having learners *experience the doing of a real task*. Another central component is to have people learn *together* under the guidance of an *expert*; but not because that's the best way to succeed in accomplishing the task -- it's because it's the best way for people to learn! Certainly, as Resnick (1987) suggests, these are two sides of the same coin -- the coin of work.

People who are learning to do something are not efficient enough to "get the job done" in time for competitive markets. People who are working, who do not have "learning support" at their fingertips, cannot learn fast enough to do the task better until, perhaps, the next time around. And people in earlier stages of physical and/or intellectual development are not up to the demanding environments of adult work.

In our view, the most advantageous aspect of fields like GDSS and CSCW is that they are concerned with supporting people, especially groups, in accomplishing tasks. The most serious deficiencies are that they are not concerned with individual or group learning, with the stages of individual development, or with promoting reflexivity about learning as their primary mission. For this reason, although the process tools of decision support and education will be similar to a degree, they are necessarily different in some respects.

Design Implications for Process Tools

To create situated environments, that is, environments which afford opportunities to engage in authentic tasks, we need to provide the same components provided in the work world. We have found a useful framework for understanding the dimensions of work, i.e., authentic tasks. Moran & Anderson (1990) describe the three aspects of what they call the *workaday world* which they present as a CSCW design paradigm. The three components are 1) technology (e.g., tools for communication, computation, composition, analysis, presentation, and so on), 2) sociality (e.g., opportunities to form social relationships both formal and informal), and 3) work practice (e.g., the knowledge, skills and routines for accomplishing specific tasks). The processes of the three are not distinct, there is a dialectic between them, and they cannot be entirely separated from each other. Thus we need to deal with constructing and studying whole environments, not just the technology we inject into them. Furthermore, the technology, i.e. tools, is there to support and "enhance, sustain, facilitate, encourage, etc., people in their work as well as be a resource for creative deployment." (Moran & Anderson ,1990, p. 387).

Tools that we create or provide must support the work practices and the sociality of the environment in authentic ways. For example, tools must provide access to source information (not just textbooks -- they don't provide the complexity or depth required) and artifacts. In order to facilitate their learning, learners must be involved actively by employing specific tools which support their analysis, personal construction and reconstruction (i.e. synthesis) of the information and artifacts (Scardamalia & Bereiter, 1985).

In studying hypertext systems, we initially saw three implications for design. First, the learner must have the ability to extract from (e.g., copy, highlight) and link to (e.g., bookmark) a content database. Second, the learner must have the ability to filter, re-word, and paraphrase from an existing database, in effect reconstructing their own version of the content domain. Third, the learner must be able to build a personal or community database in reaction to given information (e.g., commentary and critique) or from scratch (e.g., generating and explicating self-generated ideas) (Duffy & Knuth, 1990).

When considering the whole learning environment from the constructivist viewpoint we recognized the need to develop tools to support explicitly the construction by learners of their own representations and understandings. We must also build tools and strategies into the learning environment to provide authentic relationships between peers whereby they can readily communicate with each other and tap the relevant individual skills, experiences and perspectives of others in their workgroup. Similarly, our environments must support an authentic relationship between learners and experts. Thus learners must be supported in communicative and dialectical tasks with other learners and with experts as well as with the content and artifacts.

An Instantiation of Process Tools

As one part of the Enhanced Learning and Information Environments (ELIE) project (a joint research and development project of Indiana University, AT&T University of Sales Excellence, and AT&T Bell Laboratories) we have rapid-prototyped (e.g., Trip & Bichelmeyer, 1990), using HyperCard™ and Spinnaker Plus™, a networked electronic environment referred to as *RoundTable* to test our notions of process tools. In its current state of development, the *RoundTable* environment attempts to support the following processes in a social environment: comprehension, idea generation, analysis, composition, reflection, and communication.

Comprehension

To facilitate comprehension of the database we provide functions included in typical hypertext applications: note-taking, bookmarking, extraction (copying and pasting), searching, indexing, and dynamic linking. In addition, it is possible for individuals to share with others the bookmarks and links that they make.

Idea Generation

To facilitate the generation of ideas for topic-focused discussion we have developed a group brainstorming tool in which the discussion becomes part of a community database. The discussants establish new or use existing topics and sub-topics, articulate ideas and share them with the group, and react by commenting on other peoples' contributions.

Analysis

To facilitate the analysis of ideas for topic-focused discussion we have developed an analyzer tool in which the analysis becomes part of a community database. The participants establish new or use existing topics and sub-topics, articulate positions, classify positions according to a teacher-chosen logic or classification structure, share with the group, and react by commenting on other peoples' contributions.

Composition

To facilitate the construction of reports, presentations, etc., we have developed individual and group paper-writing tools where writers can create working drafts, "publish" versions to receive reactions, and then view those reactions.

Reflection

To help learners develop an awareness of their cognitive processes and development we have provided an electronic journal tool in which they are encouraged to reflect on class topics, tasks, learning strategies, group strategies, teaching strategies, and so on, as well as on the electronic environment. This journal is a private space that may be shared with the instructor.

Communication

To facilitate the self-management and coordination of group activites as well as provide the means for informal, social communication, we have provided messaging tools in which individuals can send electronic mail to other individuals including the instructor, to their work-groups, or to the entire class.

Though many of these types of tools are available commercially, we have felt it necessary to build each of them because of the need for 1) a consistent and appropriate user interface; 2) integration of information across tools; and 3) the tools to work in a unified, collaborative environment.

Case Study: Supporting Argument Analysis with RoundTable

Our first efforts in rapid-prototyping *RoundTable* involved supporting the process of argument analysis in a class "Critical Reading in the Content Areas" taught at Indiana University in the School of Education by Sharon Pugh (1990). An initial networked version of *RoundTable* was quickly developed that at the time included only comprehension, idea generation, and analysis tools. *RoundTable* was used by students synchronously (i.e., all students used *RoundTable* at the same time) to share their ideas from multiple perspectives. The students went to a computer cluster and were divided into small groups, consisting of three to four students in a group. The individuals each had their own computers and small-group members did not necessarily sit near one another. The task that we attempted to support was the analysis of case study materials portraying different viewpoint towards grading in a high school situation.

Students first started the Macintosh computer, copied the software from the server to their workstations, started the program, entered their name, selected their group, and entered an individual one-letter code (see *Illustration 1a*).

After reading through the case on-line (students were instructed to be familiar with the case materials before coming to class) and bookmarking critical passages in the text, the students used a brainstorming tool (see *Illustration 2a*) to exchange personal opinions on the case situation and characters. The brainstorm tool allowed students to react to the issues in the case as they identified them and potentially to project themselves into a similar situation as beginning teachers.

Second, an argument analysis tool (see *Illustration 3a*) provided a three part structure consisting of premises, conclusions, and evidence with which to classify the positions taken by the characters in the case. It was the instructor's perception that the argument analysis tool changed the nature of the discussion, allowing the students to focus more clearly on the task of constructing perspectives rather than interpreting issues, which appeared in the brainstorming function (Pugh, 1990).

The class used *RoundTable* for approximately one hour per day. The first day was taken up mostly by orienting the students to the Macintosh interface (most of the students had little or no prior experience with computers in general and none with the Macintosh) and having them explore and mark the case materials on-line. The students were quite verbal with questions concerning such issues as moving the mouse, clicking and double-clicking, and highlighting text.

On the second and third days the students were oriented more to the task and had substantially fewer questions and problems. They were asked at the end of class to write on a sheet of paper the best and the worst experience for each day. Students' *best* comments included: "I really enjoyed being able to immediately comment on the other group members' writing", "I like the way it works that we don't have to type the commands by ourselves but we just choose instead", "For the first time computers were fun, not frustrating. It easier sometimes to communicate on computer rather than verbally", "I'm not sure yet", "I actually remembered how to do a few commands. This gives me a slight feeling of power over my computer", "I got into the system – almost – on my own. I'm starting to understand what I am doing", "It was often easier to generate my own ideas when I was able to readily see my classmates' responses. Working with this type of computer system is efficient and fun", and "I was able to take the problem into my own hands and develop and think of my own problem-solving for the situation. It makes me feel as if my opinion is the most important."

Students' *worst* comments included: "When working a computer system for the first time its always confusing and frustrating. It is so easy to get behind when following instructions", "I'm unsure if I would feel comfortable using this system without assistance", "[The worst was] the mouse, but I'm getting better", "At this point I would prefer small group discussion", "I didn't have a worst thing, my failures were yesterday", "There is not a way to comment on a comment", "I still can't do it all myself. I'm afraid when this week is over I'll be lost again.", and "There are so many little things to remember (when to single or double-click, when to hit the escape or quit to get out of something,when putting something in the trash ejects or erases)."

The comments from the students as well as from our observing them use the tools indicated to us that 1) because of their novice level we needed to, if possible, insulate them from the Macintosh file system interface; 2) we needed to greatly automate the startup and login steps involved in gaining access to the tool; 3) from the very start, training on the tools should focus on actual and not practice tasks; and 4) care must be taken in design to indicate to the user through appropriate interface cues which actions are currently appropriate and those which are not. In general, the majority of the problems that people experienced were more concerned with the physical operation of the computer rather than the use of the tools to support the group task.

Dr. Pugh had regularly taught this critical reading technique to her class prior to our development of *RoundTable*. Our hope was to use technology to augment and extend this process. We were also curious about the effect of computer-mediated discussion on the social aspect. Research in GDSS had suggested that there are distinct differences in the exchange activities of computer-mediated and face-to-face groups (DeSanctis & Gallupe, 1987).

By the fourth day, the students were able to work independently. They directed most of the problems that they did have to other students and not to us, or used the brief "cheat sheets" that we had developed during the course of the week based on student's questions. The students, at the end of four days of using the *RoundTable* synchronously, were asked to comment – using the brainstorming tool – on their general impressions of using the environment. The full text of these evaluations can be found in the appendix. Below is an excerpt from these evaluations which we feel points to the potential benefits of using computer technologies to support mediated collaboration in the accomplishment of tasks:

> [...] It has definitely made me think more critically because sometimes when talking in small groups, it is hard to get your opinion stated either because someone else may say what you wanted to say or your thoughts just get stirred up in your head and you cannot verbally say what you are thinking. It is much easier to write down my thoughts because I can type as I think. I don't have to wait my turn to talk. It is so beneficial to see everyone else's opinions on the subject. It helps to have it right their in front of you because I can always go back and refer to someone else's statement as well as my own [...]

It was not possible to utilize RoundTable for the entire course or outside of synchronous group sessions because of the limited amount of Macintosh computers available on campus and networking obstacles that proved frustrating for both students and developers. We believe *RoundTable* provided powerful conceptual tools for students to utilize actively in an electronically supported collaborative setting. However, *RoundTable* is still in an experimental stage because of the limited networking capability on campus. In our view the instructor and students should be able to access *RoundTable* from any place on campus as well as from home, providing a powerful tool for collaborative learning. We are continuing development along these lines, working with other courses and settings at Indiana University.

Conclusions and Future Directions

It has been suggested that effective instructional design is possible only when the developers start from a theoretical basis for learning.(Bednar, et al., 1991). It was from our experience in attempting to support a specific process with technology that we have become aware of the possibilities of designing learning environments *from a constructivist epistemology*. Our experiences have emphasized for us that collaboration is not a strategy but a fundamental component of both learning and work. We are finding that capturing and evaluating the processes that these tools support is difficult and are working to develop appropriate frameworks and metrics to inform our iterative design process as well as suggesting other classes of process tools. We are interested in a variety of data, including the emerging patterns of activity by tool users as well as the quality of their constructed products, the efficiency and effectiveness of the interface design, and the affective reactions of users towards both the tools and the processes.

We realize that we need additional experience in attempting to support other types of processes if we are to envision what a full system would entail. The attempt to support argument analysis was our first attempt of constructing a *RoundTable* environment. The interaction with the students during the four days described above, as well as additional sessions with other groups of students, has led to successive iterations of and extensions to the original tools (see *Illustrations 1b, 2b, and 3b* for representations of the current interface), as well as the development of additional tools to support critical writing by a group. We believe that the best approach to defining the attributes of authentic environments for a range of authentic tasks is by observing students in their attempts to use tools in the completion of tasks and to change the tools based on student criticisms, requests, and needs.

What is clear to us is that starting from the basis of constructivist epistemology leads to radically different approaches to the design of learning environments and the need to glean from other fields directions for conceptualizing, designing, and making sense of processes of work and learning.

References

Bednar, A. K., Cunningham, D., Duffy, T. M., Perry, J. D. (1991). Theory into practice: How do we link. In G. J. Anglin (Ed.), *Instructional technology: Past, present, and future.* Englewood, CO: Libraries Unlimited.

Brown, A. L. (1981). Metacognition: The development of selective attention strategies for learning from texts. In M. L. Kamil (Ed.), *Directions in reading: Research and instruction.* Washington, DC: National Reading Conference.

Brown, J. S. (1989). Toward a new epistemology for learning. In C. Frasson & J. Gauthiar (Eds.), *Intelligent tutoring systems at the crossroad of AI and education.* Norwood, NJ: Ablax.

Brown, J. S., Collins, A. and Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*, 32-42.

Byers, T. J. (1987, April). Built by Association. *PC World*, 245-251.

Cobb, P., & Steffe, L. P., (1983). The constructivist researcher as teacher and model builder. *Journal for Research in Mathematics Education, 14*(2), 83-94.

Cunningham, D. (1987). Outline of an educational semiotic. *The American Journal of Semiotics, 5*, 201-216.

DeSanctis, G., & Gallupe, R. B. (1987). A foundation for the study of group decision support systems. *Management Science*, 33(5), 589-609.

Duffy, T. M. & Knuth, R. A. (1990). Hypermedia and instruction: Where's the match? In D. Jonassen (Ed.), *Designing Hypermedia for Learning.* NY: Springer-Verlag.

Garner, R. (1987). *Metacognition and reading comprehension.* Norwood, NJ: Ablex Publishing Corporation, 105-126.

Greif, I. (Ed.) (1988). *Computer-supported collaborative work.* Los Altos, CA: Morgan Kaufman.

Hammond, N. (1989). Hypermedia and learning: who guides whom? *ICCAL Proceedings '89*, 167-181.

Kraut, R. E. (1989). Telecommuting: The trade-offs of home work. *Journal of Communication, 39*(3), Summer 1989, 19-47.

Locatis, C., Letourneau, G., & Banvard, R. (1989). Hypermedia and instruction. *Educational Technology Research and Development, 37* (4), 65-77.

Marchionini, G. (1988, November). Hypermedia and learning: Freedom and chaos. *Educational Technology*, 8-12.

Merrill, M. D. (1983). Component Display Theory. In C. M. Reigeluth (ed.), *Instructional-Design Theories and Models: An Overview of their Current Status*. Hillsdale, NJ: Erlbaum.

Moran, T. P., & Anderson, R. J. (1990). The workaday world as a paradigm for CSCW design. Proceedings of the Conference on Computer-Supported Cooperative Work, October, 1990. Los Angeles, CA.

Newman, D. (1990). Opportunities for research on the organizational impact of school computers. *Educational Researcher, 19*(3), 8-13.

Pugh, S. (1990). Examining student change in a teacher education course using dialectical reasoning as a strategy for critical reading. Paper presented at the National Reading Conference, Miami, FL.

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher, 16*(9), 13-20.

Resnick, L. B. (1985). Instructional psychology. In T. Husen, & T. N. Postelthwaite (Eds.), *The international encyclopedia of education*, 2569-2581.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions*. Hillsdale, NJ: Erlbaum.

Streibel, M. J. (1986). A critical analysis of the use of computers in education. *ECTJ, 34*(3), 137-161.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Appendix: Student Evaluations of RoundTable

The tool has been helpful to see many perspectives on a case. I really benefitted from seeing my classmates comments. I like this tool although it was a little tough at first. It can be helpful and would be ok to use for the semester to see others comments. Once you learn the commands it becomes much easier and not really discouraging. I enjoyed this experience and am glad I had this opportunity. The tool is great. RT is better than EC!!!!!!!! [1] This was a very beneficial learning experience. I wish that we had more time to work with the program. I think that it would be great if future X401 students could use this to analyze cases.

Once I got the hang of working the system, it was really easy to generate ideas concerning the case. The system enabled me to quickly jot down my own opinions and then I was able to look at the opinions contributed by other members in my group. Additionally, when I was stuck on a particular topic, for example Frank's warrants, I could simply go to another topic and work on that one until I had come up with some ideas for the previous topic.

This program is great. It has definitely made me think more critically because sometimes when talking in small groups, it is hard to get your opinion stated either because someone else may say what you wanted to say or your thoughts just get stirred up in your head and you cannot verbally say what you are thinking. It is much easier to write down my thoughts because I can type as I think. I don't have to wait my turn to talk. It is so beneficial to see everyone else's opinions on the subject. It helps to have it right their in front of you because I can always go back and refer to someone else's statement as well as my own. All in all, this was a valuable experience. For the first time the computer was fun and I looked forward to coming each day.

I am so glad that I have been able to use this program and feel very satisfied with the fact that I have learned to use the whole program in four days. Except for the kinks, everything is wonderful. When can we use this in our classrooms? How can I make further use of this program and how do you program in your own lessons? Is this an experimental program? Round Table would be excellent to use in classrooms and allows critical thinking and commenting with out limiting it to discussion times within the classroom. Every student can comment on a topic and interact with their peers in quality time span. Great! Thanks for the help and patience. Better than EC.

I think that the RT is a good tool to use and to develop critical thinking skills. Reading people's perspectives can increase your own. I have read many perspectives that I never would have thought of. The CHEAT SHEET is a tremendous help. After spending the week in this lab, I think my skills for these MACS are increasing (still a long way to go). Comparing the RT to the EC, the RT is better because the thoughts are organized and you don't have to go through everybody's thought to get to the one you want to comment about. The RT seems better structured. It seems that with the RT, there is more to comment about. I am not sure if it because it is divided up or what!

I see how this type of activity can be useful if, for no other reason, it is different and, therefore, interesting. It could break the monotony of the normal classroom. I like the opportunity to be able to share ideas all at once without having to interrupt one another. The argument analyzer has been positive in that it has really encouraged dialectical thinking. We've had to take each person in the "Making the Grade" and look at the same situation from their varied perspectives. The most positive thing to me has been another opportunity to get my hands on a computer and grow to be more comfortable in using it. Thanks!

[The] cheat sheet would have been more helpful on first or second day rather than last. I like the argument analyzer, the set up helps organize your thoughts and shows your argument clearly and that of others in your group . I would have liked to see opinions of other group their perspective is different.

---

[1] Electronic Classroom -- a menu-driven, character-based topic oriented discussion tool on the VAX computer.

I think this was probably a good experience even though I don't like to use computers. I found the Round Table program interesting. I think that you underestimate the amount of work that could be done in an hour. I found myself getting bored because I was finished with what we were supposed to be doing. I also felt lost most of the time as to what was too be completed. I think the class could be more productive if we are told specifically what needs to be done or what we can work on. Sometimes I wanted to just write, like a journal entry, but I wasn't sure if I was allowed to do that, and I didn't know where to do that. I think this program would be good to use in an English that I want to teach on ethics. I think it would help students to open up and feel free to write down what they wanted without feeling that they would be criticized. I also like the idea that the students would be able to write back and forth with comments.

give me a week. I'd love to learn this system. I like the fact that it can hook up to its resources at almost any given moment, like the link hook up. I need to learn how to shuffle the cards better. And the filing is a little hard to understand at times, but I suspect that the reason is because I Am not familiar with all of the language yet.

I think the cheat sheet was very helpful.

I really liked using this system. I like the graphics. The best part about this program was that we were able to view the ideas of everyone in our group and then comment on them. Reading other reactions helped to form my own thoughts. Thanks for teaching us how to use this program. I didn't like not being able edit comments that I've entered. I also wish that I could've seen student opinions other than those in my group. There were only 3 of us, so we had a limited number of statements.

There have been times that I wanted one of my group members to read a response that I made to one of their comments. But there was no way to call their attention to it. I just had to hope that they would eventually update in that particular window. (Or I could walk over to them and call their attention to it). Overall, this has been an interesting way to interact with class members. I would not mind using it once in awhile, but my preferred method would be face to face. The more I use it the more comfortable I will become and may wish to use it more.

In my opinion, this computer program is very useful. It helps a lot in making doing the assignments easier and more convenient. Although it may seem to be quite complicated at first, it is not too difficult once we get used to it. All the commands are also not too difficult to remember. I like the idea that we can work together with other people in the same group. For example, we have a chance to exchange our ideas and opinion. Once again, it is really useful and I like it.

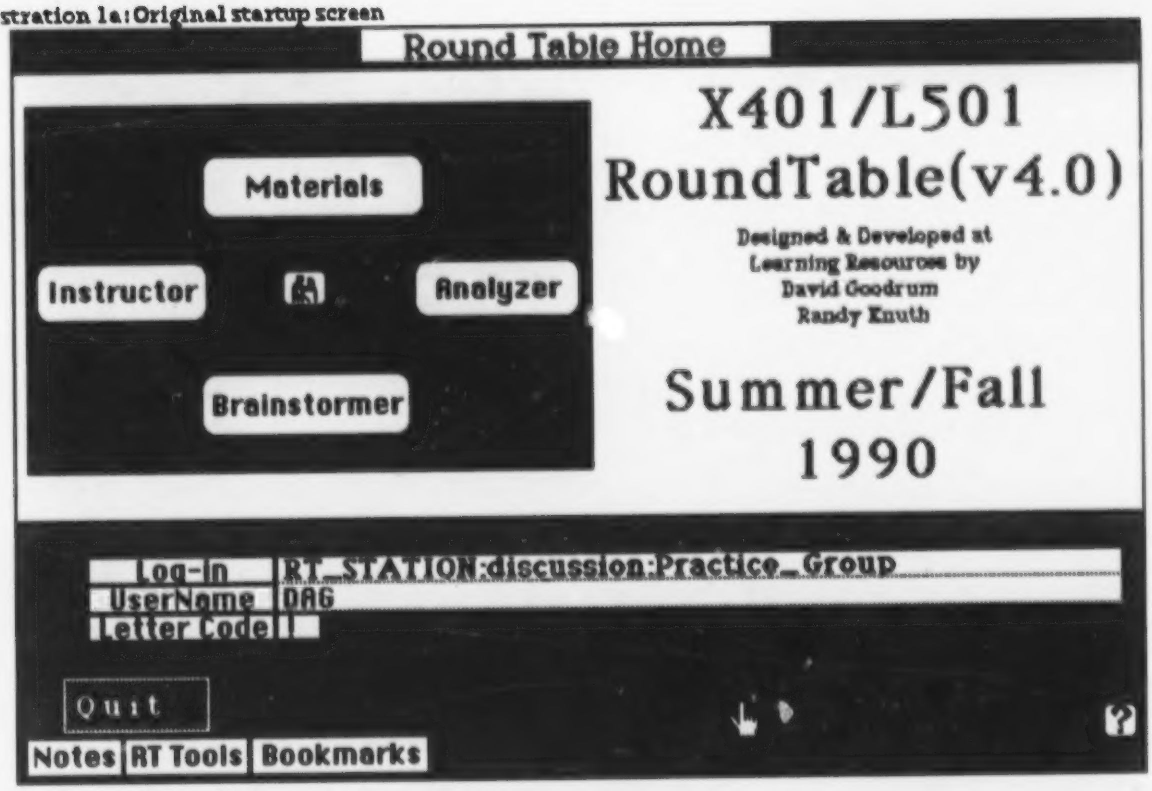
Illustrations

On the following pages are sample screens of tools in *RoundTable* at two distinct points in time. The first of each set is the original version used with Sharon Pugh's class. The second of each represents the same tool at the current time.

One of the guidelines we used for the first iteration was that the capabilities of the tool should be explicitely available to the user. Thus, buttons were provided for the major functions of the tool. Based on reactions of students who used the tool over a period of four days, we felt that this approach basically worked, but that the screen caused some confusion because it was unclear which button was the *appropriate* one at any given point in time. In the current design, buttons are not displayed on the screen unless they are part of the current action. For example, the "Add" button is only displayed when the person is in the process of writing a statement and then clicking the button signals that the person is finished and wishes to share the statement with the group. The "Comment" button is only displayed while the person is making a comment on a specific statement. The major functions are still made visibly available to the user through the pull down menus at the top. When a person pulls down a menu, only the appropriate choices for the current task are displayed in bold type; the others are shaded grey.

The most significant addition as a result of *RoundTable's* use in Sharon Pugh's class, was the addition of mail, initially at the group level. Since the sessions were synchronous, and their discussion was updated almost continuously, it had not occured to us that they would have trouble communicating with each other. But there was a specific expressed need for a way to coordinate group direction.

Illustration 1a: Original startup screen



Round Table Home
Materials
Instructor
Analyzer
Brainstormer
Quit
Notes
RT Tools
Bookmarks

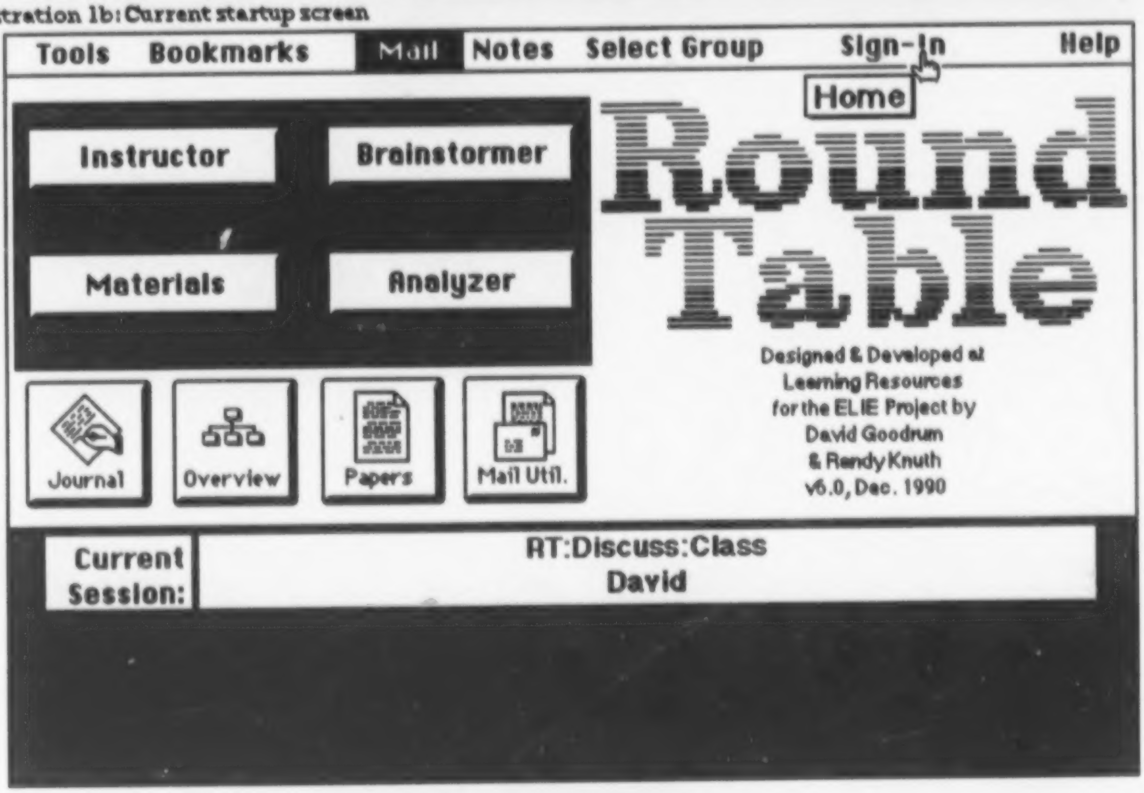
X401/L501
RoundTable(v4.0)

Designed & Developed at
Learning Resources by
David Goodrum
Randy Knuth

Summer/Fall
1990

Log-In	RT_STATION:discussion:Practice_Group
UserName	DAG
Letter Code	!

Illustration 1b: Current startup screen



Tools
Bookmarks
Mail
Notes
Select Group
Sign-In
Help
Home
Instructor
Brainstormer
Materials
Analyzer
Journal
Overview
Papers
Mail Util.

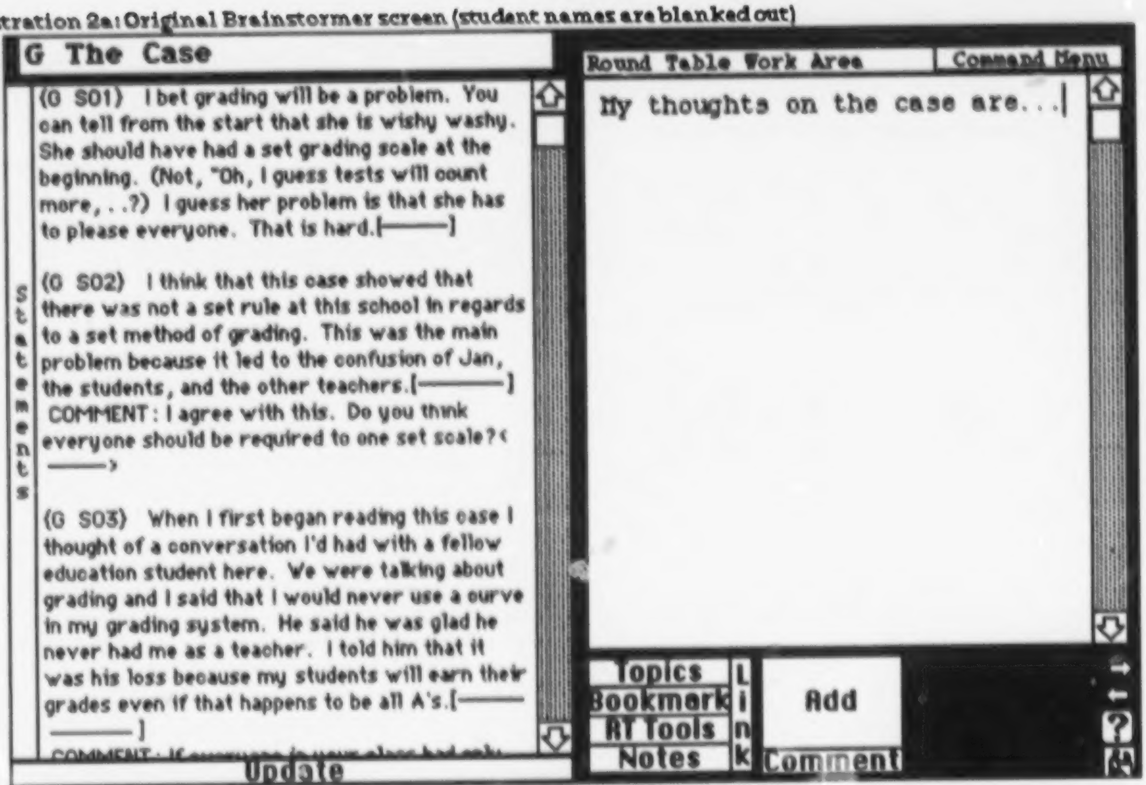
Round
Table

Designed & Developed at
Learning Resources
for the ELIE Project by
David Goodrum
& Randy Knuth
v6.0, Dec. 1990

Current Session:	RT:Discuss:Class David

BEST COPY AVAILABLE

Illustration 2a: Original Brainstormer screen (student names are blanked out)



(G S01) I bet grading will be a problem. You can tell from the start that she is wishy washy. She should have had a set grading scale at the beginning. (Not, "Oh, I guess tests will count more, . .?) I guess her problem is that she has to please everyone. That is hard.[———]

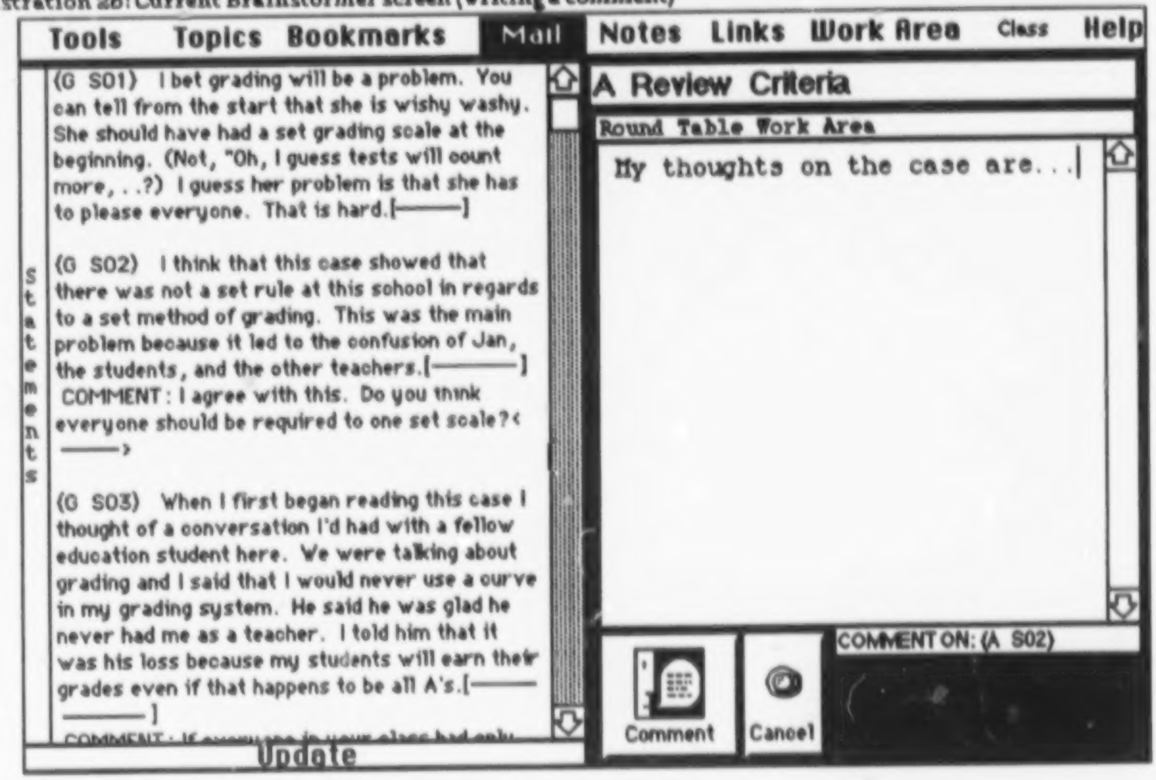
(G S02) I think that this case showed that there was not a set rule at this school in regards to a set method of grading. This was the main problem because it led to the confusion of Jan, the students, and the other teachers.[———]
COMMENT: I agree with this. Do you think everyone should be required to one set scale?< ———>

(G S03) When I first began reading this case I thought of a conversation I'd had with a fellow education student here. We were talking about grading and I said that I would never use a curve in my grading system. He said he was glad he never had me as a teacher. I told him that it was his loss because my students will earn their grades even if that happens to be all A's.[——— ———]

My thoughts on the case are...

G The Case
Statements
Round Table Work Area
Command Menu
Topics
Bookmark
RT Tools
Notes
Link
Add
Comment
?
Update

Illustration 2b: Current Brainstormer screen (writing a comment)



(G S01) I bet grading will be a problem. You can tell from the start that she is wishy washy. She should have had a set grading scale at the beginning. (Not, "Oh, I guess tests will count more, . .?) I guess her problem is that she has to please everyone. That is hard.[———]

(G S02) I think that this case showed that there was not a set rule at this school in regards to a set method of grading. This was the main problem because it led to the confusion of Jan, the students, and the other teachers.[———]
COMMENT: I agree with this. Do you think everyone should be required to one set scale?< ———>

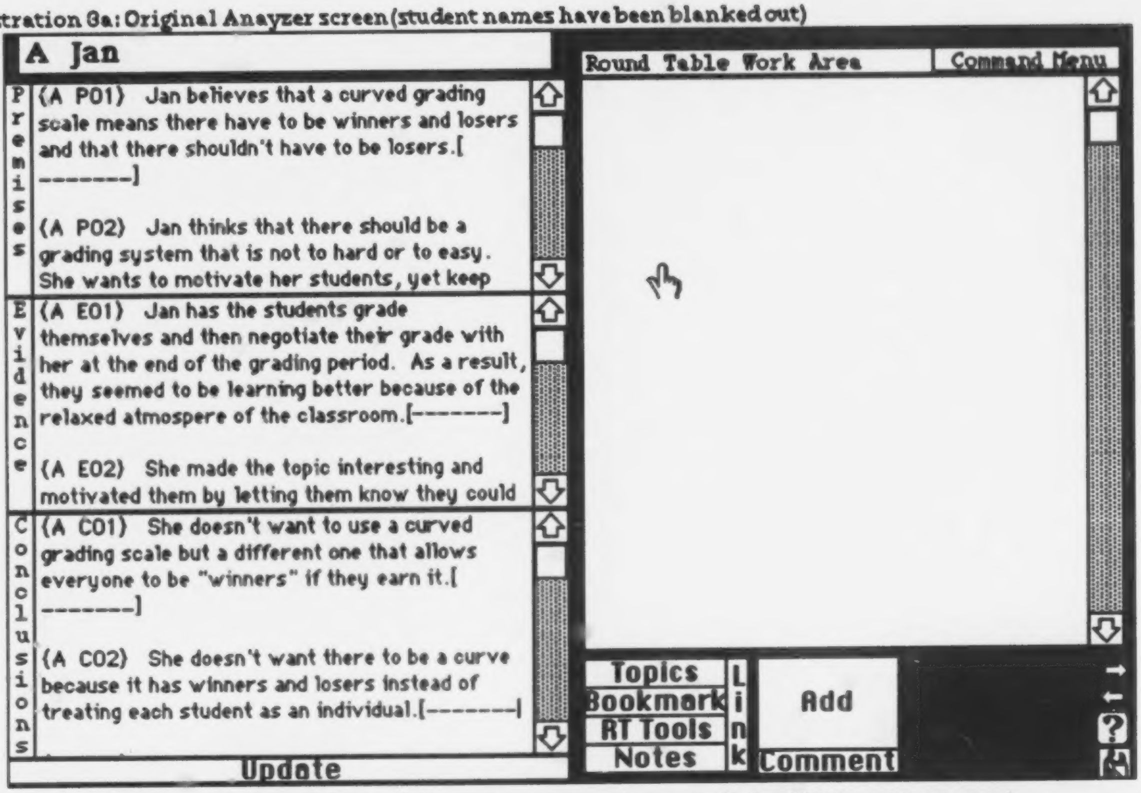
(G S03) When I first began reading this case I thought of a conversation I'd had with a fellow education student here. We were talking about grading and I said that I would never use a curve in my grading system. He said he was glad he never had me as a teacher. I told him that it was his loss because my students will earn their grades even if that happens to be all A's.[——— ———]

My thoughts on the case are...

COMMENT ON: (A S02)

Tools
Topics
Bookmarks
Mail
Notes
Links
Work Area
Class
Help
Statements
A Review Criteria
Round Table Work Area
Comment
Cancel
Update

Illustration 8a: Original Anayzer screen (student names have been blanked out)


A Jan
Premises
Evidence
Conclusions
Update
Round Table Work Area
Command Menu
Topics
Bookmark
RT Tools
Notes
Link
Add
Comment
?

(A P01) Jan believes that a curved grading scale means there have to be winners and losers and that there shouldn't have to be losers.[-------]

(A P02) Jan thinks that there should be a grading system that is not to hard or to easy. She wants to motivate her students, yet keep

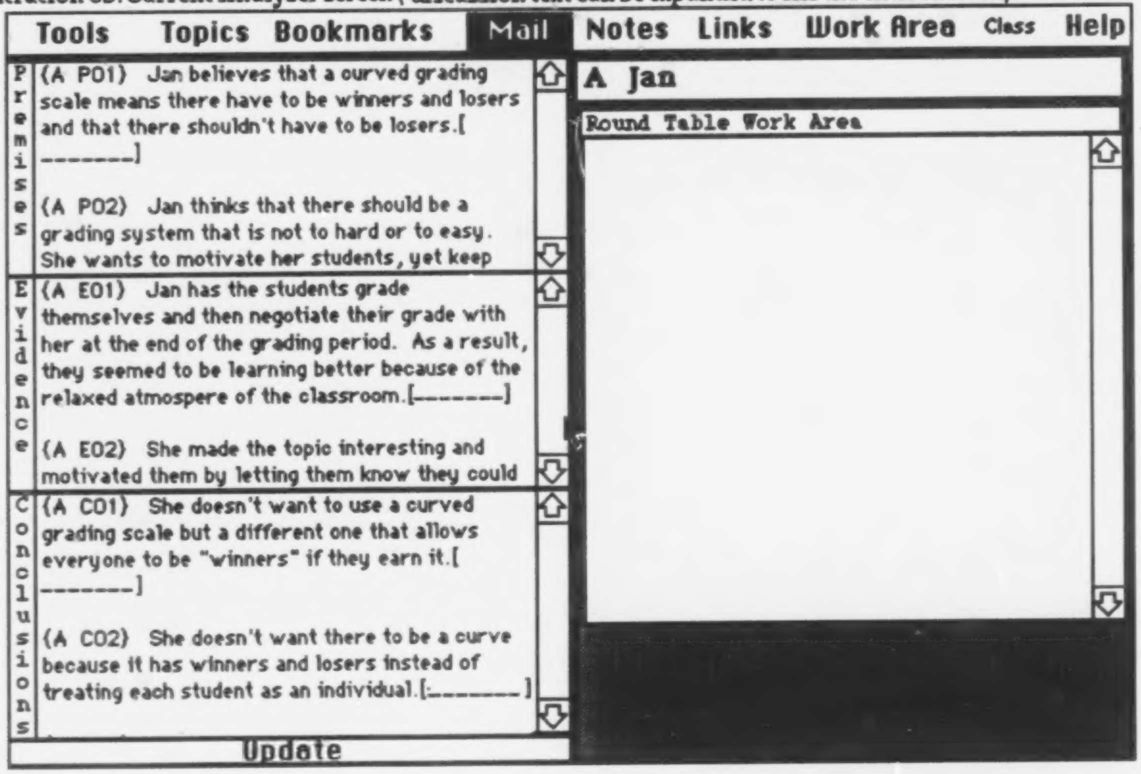
(A E01) Jan has the students grade themselves and then negotiate their grade with her at the end of the grading period. As a result, they seemed to be learning better because of the relaxed atmospere of the classroom.[-------]

(A E02) She made the topic interesting and motivated them by letting them know they could

(A C01) She doesn't want to use a curved grading scale but a different one that allows everyone to be "winners" if they earn it.[-------]

(A C02) She doesn't want there to be a curve because it has winners and losers instead of treating each student as an individual.[-------]

Illustration 8b: Current Analyzer screen (discussion text can be expanded to fill the entire screen)


Tools
Topics
Bookmarks
Mail
Notes
Links
Work Area
Class
Help
Premises
Evidence
Conclusions
Update
A Jan
Round Table Work Area

(A P01) Jan believes that a curved grading scale means there have to be winners and losers and that there shouldn't have to be losers.[-------]

(A P02) Jan thinks that there should be a grading system that is not to hard or to easy. She wants to motivate her students, yet keep

(A E01) Jan has the students grade themselves and then negotiate their grade with her at the end of the grading period. As a result, they seemed to be learning better because of the relaxed atmospere of the classroom.[-------]

(A E02) She made the topic interesting and motivated them by letting them know they could

(A C01) She doesn't want to use a curved grading scale but a different one that allows everyone to be "winners" if they earn it.[-------]

(A C02) She doesn't want there to be a curve because it has winners and losers instead of treating each student as an individual.[-------]

Title:

CBI: Systems or Medium?

Author:

Nancy L. Higginbotham-Wheat

CBI: System or Medium?

Abstract

Nancy L. Higginbotham-Wheat
Memphis State University

This paper addresses one area of conflict in decision-making in CBI research: The relationship between the researcher's definition of computer-based instruction (either as a medium or as an integrated system) and the design of meaningful research questions. Through a review of the literature, this paper not only explores the differing definitions of CBI as either a system or a medium but provides examples of CBI utilized as each. Implications for researchers based on the definition of CBI as a system or a medium are discussed, and guidelines in evaluating current research in the field are provided. A list of 31 references is included.

Keywords: computer-based instruction, instructional design, instructional technology, research design, media research.

CBI: System or Medium?

Nancy L. Higginbotham-Wheat
Memphis State University

In 1986 it was estimated that there were more than one million microcomputers in use in elementary and secondary schools in the United States and that this number would increase to approximately 2.8 million in the near future (National Task Force on Educational Technology, 1986). As a result of this microcomputer proliferation in schools there has been a corresponding increase in the quantity of research studying the educational uses and effects on learning of computer-based instruction (CBI). It has been difficult to discern among this wealth of information just what questions constitute meaningful research and how best to approach the research itself. Researchers interested in computer-based instruction come from a variety of disciplines, from instructional design and media instruction to educational psychology and computer science, and, thus, bring different perspectives--indeed, even differing levels of expertise in research methodology to the study of computer-based instruction. The research is still at an adolescent stage of development, with the concomitant conflicts in goal development through which any discipline (or individual for that matter) is obliged to make passage.

One particular area of conflict has been the use of media comparison studies in the study of CBI. According to Salomon and Clark (1977), much of the early research reflects "research with media" as opposed to "research on media". When the research design focuses on the content or methods carried by a medium but not the medium itself, the study represents research with media. For example, to investigate the effects of summary lists on the learning of new information, the textual material and the summary lists could be presented on a computer screen. Discovering that the lists improved students' scores on posttests tells us something about the use of summaries but nothing about the nature of the computer as an instructional message vehicle. If we could compare the results of this study on the computer with the same strategy in textbook format we would expect the same results. Since the same instructional strategy is employed, it requires similar behavior on the part of the learner.

In the relatively uncomplicated study above it is a simple matter to separate media and methods, but in many studies we (or our readers) are unable to do so completely (Hagler & Knowlton, 1987). With highly interactive technologies, media and methods are confounded. For example, interactive video allows immediate review from any point in the lesson. Many computer programs can respond to students' on-task input with corrections and/or prescriptions and remediation. Although less interactive media could incorporate some of these strategies, many of these strategies are unique to the technologies of instruction for which they were designed and cannot "carry over" to a variety of media. Media researchers are familiar with Clark's (1983) analogy comparing media to delivery trucks and instructional content to the groceries carried by the truck. Clark's contention is that the choice of the truck (medium) is unimportant from a nutritional (achievement) point of view and that, where differences are found in achievement, it is the instructional method used that is responsible for the "superiority" of one medium over another. From this perspective, what is gained from media research is a theoretical understanding of learning in general that is applicable over a wide variety of media. When CBI research is evaluated, therefore, one must determine whether investigators regarded CBI as a medium or as an integrated delivery system.

At this point, the reader could justifiably ask the meaning of these two terms, medium and system. Knowlton (1964) points out that there are two types of definitions, real and nominal. If a "thing" is perceived in much the same way by members of a language group then to define it is but to name it, hence the term "nominal" definition. On the other hand, if a "thing" is perceived differently and understood differently by several members of a language group, giving it a name will not facilitate understanding. A single term for a whole category of widely varying concepts sometimes causes people to believe they are speaking of the same "thing" when this is not the case at all. Knowlton (1964) states that disagreements among members of the same discipline in regard to key terms of the discipline are quite common until the discipline "matures". Such is the state of affairs in our adolescent discipline of computer-based instruction. Real definitions make reference to the criterial attributes of the thing defined, implying some sort of conceptual scheme relevant to our interests. The aim is to achieve a conceptual mechanism to help the researcher avoid the trivial and to ask the incisive question. Salomon (1970) defines an instructional medium as a package of unique modes of presenting information which also fulfill a unique psychological function. According to his definition, if a picture in a book and on a screen fulfill the same psychological function they are one instructional medium (given common learners). Another way Salomon (1970) defines medium is as the overlapping area of two circles: the stimulus-attributes circle and the response circle. When various stimuli have common structural attributes but do not call for common cognitive responses, they cannot be said to constitute one medium.

Media were taxonomized according to stimulus attributes for a number of years in instructional technology. One of those who left this approach of classification was Heidt (cited in Cassidy, 1982), who argued for a taxonomy based on learner responses rather than on stimulus characteristics. Heinich, Molenda, & Russell (1982) define 'medium' in a general sense as a means of communication referring to anything that carries information between a source and a receiver.

Once we have arrived at a working definition of the term 'medium', we have to explore the term 'system' for contrast. Romiszowski (1981) contends that the traditional definitions, such as "a set of components or elements, interacting together towards a common goal", do not provide a fruitful approach. He uses the analogy of a bicycle to illustrate the point that a system exists because someone has defined it as such. The bicycle has clearly defined components--wheels, axle, pedals, chain--and an agreed upon goal or purpose (propelled transport). One might consider the bicycle,however, as only a part of a larger system of urban transport or as a carrier of a smaller system, the gears of the back wheel. Thus, for practical purposes, a system exists because we have drawn the boundaries that limit the extent of the system, defining the components or subsystems that make up the 'system of interest'. We can identify the reason for the existence of the system by examining its relationship to its environment, its inputs and outputs. From this comes the concept of the 'black box', which is analyzed by quantifying the inputs and outputs and, indirectly, the efficiency of the system.

The casual reader might question the importance of these two definitions of 'system' and 'medium', but their importance to the instructional technologist cannot be overestimated. If CBI is a system, it incorporates the medium, the learner, the designer, all the experimental variables, the hardware, the learning environment, the software--all integrated and working together toward the intended outcome: learning. The medium is the delivery device of the system. The true CBI system consists of characteristic strategies that one would normally find from one program to the next, i.e., interactive responding, use of graphics, use of feedback, active learner response, branching capabilities. The computer, while it is a piece of hardware and a medium for instruction, is only part of the system of CBI, a set of components (strategies) defined by the 'system of interest'. Some researchers define the computer as Type 1 technology, or hardware, much the same way one would describe an overhead projector. Indeed, even though they they contend that, by itself,the computer is of little value (by implication supporting the systems approach) Hannafin & Peck (1988) refer to computer-assisted instruction as an educational medium throughout their entire text on design. What should make the computer interesting to the researcher, however, are not the *attributes of the hardware* but the components of the system of which it is an integral part. If CBI is a true system, there are certain types of research questions which become valid and viable for study which would not be otherwise. If CBI is a system then researchers can compare systems, say, CBI to audio-tutorial systems, without the label of "media comparison" studies. Basic research could then address questions regarding theoretical understanding of how systems work to affect learning and motivation, while applied or developmental research could go about the business of improving technology as an educational tool by testing and evaluating theoretical positions.

Clark (1983) argued that media comparison studies did not yield results that suggested media had any influence of learning under any conditions. The major problem Clark addressed in his criticism of these types of studies is the inability of researchers to control all the confounding variables associated with this type of research. Clark reported the most common sources of confounding in media research seem to be the uncontrolled effect of instructional method or content differences between treatments that are compared and a novelty effect which disappears over time. Ross & Morrison (1988) state that an even more critical concern is whether it is sensible to compare media in the first place since they merely serve as delivery devices, not as the content or the instructional strategy used to teach content. No one challenges Clark's assertion (1983) that, "It seems not to be media but variables such as instructional methods that foster learning." His statement, however, that most computerized instruction is "merely the presentation of PI or PSI via a computer" illustrates a basic misconception that the computer--the hardware-- is the area of interest, not the characteristic strategies employed for each.

Clark (1985) criticized the meta-analysis of 500 studies for classroom-based comparisons of CBI and conventional teaching conducted by Kulik and associates (Kulik, Kulik, & Cohen, 1980). The thrust of Clark's evidence for confounding in these studies was that, on the one hand, methods such as the use of examples and matched non-examples, individualization of pacing, corrective feedback after response, and a close correspondence between instruction and test items tend to be used in the design of CBI lessons but not by teachers in comparison treatments. On the other hand, another source of confounding might be that teachers compete with the computer as a perceived 'rival', thereby diminishing the true effect of CBI on student achievement. In his discussion of the meta-analysis Clark (1985) found serious design flaws in many of the studies which showed positive results with CBI. Much of the argument points toward Clark's perception that the non-CBI treatments were not 'robust'. In studying Clark's review of the meta-analysis, however, this researcher read selected studies from the subset of Clark's reviewed studies and found clear evidence that what was reported as CBI in many studies was simply the employment of the 'hardware' in some fashion. For example, Cox (1974) reported a use of computer-assisted instruction in the use of computer 'simulations' and 'games'. Upon reading the methodology, however, it becomes clear that this was a misuse of the term 'computer-assisted instruction'. The instructor set up simulations of various economic principles, feeding set variables into the computer. Printouts were distributed to the students, who were then told what the printouts were supposed to illustrate. The closest the student came to the computer was handling the paper artifact of the printout.

Because this method proved ineffective in teaching economic principles, all CBI systems were predicted to be equally ineffective. While the experiment above is an example of inappropriate research labeled as CBI, there were also included in Clark's (1985) review some particularly well-designed studies employing true CBI, which, incidentally, seemed to also report a higher incidence of achievement. For example, Green & Mink (1973) employed computer assisted simulations of experiments in eyelid conditioning as part of a 2-week component of introductory psychology compared to two other groups: one group receiving only traditional instructor-led seminars and another group receiving seminars with computer-generated information but not delivered by computer. While none of the groups were able to generalize principles of scientific analysis to other problems, the computer-assisted group performed better on the tasks of labeling variables and controls. The focus of the study by Jones & Sorlie (1976), included in Clark's (1985) review as an experiment utilizing CBI, was on computer-assisted testing. It nevertheless contained information of interest to the designer of CBI. Medical students studying for comprehensive board exams were separated into two groups, one of which had paper and pen testing only and the other with CAT (computer-assisted testing). The CAT group was reported as performing significantly higher on their board exams, attributed by the authors to a cumulative effect of the review process offered by the CAT method. According to Jones & Sorlie (1976) results show that "...the advantage was not due solely to the computer-assisted medium, but to the process made possible by the medium: a 'teaching exam' with interactive review and immediate feedback."

Other studies criticized by Clark (1985) were no more than page-turning uses of the computer (Broh, 1975) or electronic worksheets (McEwen & Robinson,1976). The most common form of instructional method was some form of drill and practice, a weak use of the CBI format (Suppes & Morningstar, 1969; Chiang, Stauffer, & Cannara, 1978). Another misuse of the term CBI was found in the study of PLAN, a computer-managed program for disadvantaged youth, by Coffman & Olsen (1980). Some of the studies found significant differences favoring a control group (Axeen,1971) where the same teacher effect might have provided some sort of compensatory rivalry with the computer as teacher, thus confounding the study. Most of these studies, therefore, exhibit design flaws and/or weak uses of CBI, making any generalizations difficult to achieve concerning the efficacy of the CBI system. Clark's contention that the CBI achievement gains were overestimated due to uncontrolled and robust instructional methods embedded in CBI treatments ignores the fact that most of the studies were, indeed, weak uses of a systematic set of strategies inherent in true CBI systems design. If stronger, more effective uses of CBI had been examined, the implication arises that achievement gains would have been more impressive for the CBI groups rather than less so.

As a response to the questions raised concerning the impact media have on learning and media research designs, a continuing debate has been carried out in instructional design journals and at the annual meetings of professional organizations in the field. Simonson (1987) led a symposium on the status of research in instructional technology which brought together researchers advocating extreme positions in the debate. In the first of four papers presented in the symposium, Clark (1987) maintains his original position that media have no impact on learning. He also takes the position that media attribute studies are not supported by findings in cognitive research. He suggests that more profitable studies in media research would result from abandoning the older "craft" approach to media in favor of a newer scientifically-based technology approach and argues for a renewed commitment to and training in research methodology.

Kulik and Kulik (1987), in the second symposium paper, present results of four separate statistical analyses of findings on computer-based instruction which indicate that computer-based instruction does indeed have positive effects on students. The authors discuss three factors which might have contributed to favorable reports in the literature: editorial gatekeeping, experimental design flaws, and instructional quality. They suggest that research continue, focusing on different research designs to investigate reasons for differences.

In the third of the symposium papers, Tennyson (1987) elaborates on the basic question of the efficacy of media in improving learning. According to the author, the problem is not the technology, but , rather, the failure of proponents to adequately trace the variables of their respective media techniques to clearly defined learning processes. He argues that, if instructional methods and/or media are to improve learning, they must have a direct trace to a specific learning process and empirical support demonstrating their significance. Focusing on the tracing process, Tennyson presents an information processing model of learning as a means by which researchers may determine if specific media variables and methods may improve learning. Secondly, he maintains that media "...is but one component in a complex instructional system. A system that involves principles of instructional design as well as methods of instructional delivery" (p.8).

Winn (1987) concluded the symposium by identifying five aspects of media that might affect learning: 1) the technology used; 2) the symbol systems employed; 3) the content conveyed; 4) the setting in which media are employed; and 5) thoroughness of message design. Since Clark's (1983) argument against media studies deals only with the technology for delivery of instruction, Winn examines the other four factors might affect learning, including how these factors relate to technology.

Although the symposium is of interest to all media researchers, the debate continues over what constitutes meaningful research of media variables. Simonson's (1987) hope that the papers and discussions presented would help “clear the air” on this issue may not have been realized, but they may serve, nevertheless, as guides along the paths toward meaningful media research .

Summary

If CBI is defined as merely a medium, the nature of legitimate research becomes limited, perhaps only to those designs which ask questions concerning strategies of instruction. Failure to limit research to these legitimate questions gives rise to the media comparison studies which characterized early media research. According to Salomon and Clark (1977), much of this early research reflects "research with media" as opposed to "research on media". Unfortunately, even a cursory review of current literature reveals that many modern research designs focus on the content or methods *carried* by a medium but not the medium itself. These studies represent research with media.

If CBI is a system, it incorporates the medium, the learner, the designer, the hardware, the learning environment, the software (including the displays, user input, and program decisions). These factors are all integrated and work together toward the intended outcome: learning. *The medium becomes the delivery device of the system.* The true CBI system consists of characteristic strategies that one would normally find from one program to the next, i.e., interactive responding, use of graphics, use of feedback, active learner response, branching capabilities.

The computer, on the other hand, is a piece of hardware and a medium for instruction, only part of the system of components (strategies) defined by the "system of interest"— CBI. Some researchers (Heinich, Molenda, & Russell, 1988; Davies, 1971; Lumsdaine, 1964) define the computer as Type 1 technology, or hardware, much the same way one would describe an overhead projector. What should make the computer interesting to the researcher, however, are not the attributes of the hardware but the components of the system of which it is an integral part. If CBI is a true system, there are certain types of research questions which become valid for study which would not be otherwise. If CBI is a system then researchers can compare systems, e.g., CBI to audio-tutorial systems, without the label of "media comparison" studies. Basic research could then address questions regarding theoretical understanding of how these systems work to affect learning and motivation, while applied or developmental research could go about the business of improving technology as an educational tool by testing and evaluating theoretical positions. When CBI research is evaluated, therefore, one must determine whether investigators regarded CBI as a **medium** or as an **integrated delivery system** that ***fully*** utilizes the attributes of the computer to enhance the delivery of the instruction.

If CBI is no more than a delivery vehicle as Clark's grocery truck analogy maintains, with no effect on the analogous nutritional (learning) requirements of the consumer, then CBI is just another piece of fancy hardware with no inherent, embedded strategies at work. As Morrison, Ross, & O'Dell (1988) have asserted, however, the delivery truck is defined many times by the attributes (e.g., refrigeration units) which insure that the groceries delivered are in peak condition. By extension of the analogy of the grocery truck one can begin looking at the entire system of suppliers to the consumer. The dairy operation, for example, is not defined only by the cows in the pasture nor is it limited to the refrigerated trucks which deliver the milk. It is, however, an integrated system working together to achieve its particular goal of milk supply. One can find fault in the taste of the milk but cannot attribute that fault to the system itself, rather to a fault in one part of the system, whether it is a breakdown in the refrigeration unit or inadequate feed for the cow.

References

Berlo, David K. (1960). *The Process of Communication: An introduction to theory and practice*. New York. Holt, Rinehart and Winston.

Cassidy, Michael F. (1982). Toward integration: Educational technology and semiotics. *Educational Communication & Technology Journal, 30*,2.

Clark, Richard E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*,4.

Clark, Richard E. (1985). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. *Educational Communication & Technology Journal, 33*,4.

Clark, Richard E. (1987). *Which technology for what purpose? The state of the argument about research on learning from media*. Paper presented at the Annual Convention of the Association for Educational Communications and Technology. Atlanta, GA. ERIC Document Reproduction Service No.ED 285520.

Hannafin, Michael J., & Peck, Kyle L. (1988). *The Design, Development, and Evaluation of Instructional Software*. New York. Macmillan.

Heinich, Robert, Molenda, Michael, & Russell, James D. (1982). *Instructional Media*. New York. John Wiley & Sons.

Heinich, R. (1984). The proper study of instructional technology. *Educational Communication and Technology Journal, 32*,(2),67-87.

Knowlton, James Q. (1964). A conceptual scheme for the audiovisual field. *Bulletin of the School of Education, Indiana University,40*,3.

Kulik, J. Kulik,C., & Cohen,P. (1980). Effectiveness of computer-based college teaching: A meta-analysis of findings. *Review of Educational Research, 50*, 525-544.

Kulik, J. & Chen-Lin, C. (1987). Computer-based instruction: What 200 Evaluations Say. Paper presented at the Annual Convention of the Association for Educational Communications and Technology, Atlanta, GA. ERIC Document Reproduction Service Number ED 285521.

Morrison, G.R., Ross, S.M.,& O'Dell, J.K. (1988). Text density level as a design variable in instructional displays. *Educational Communication & Technology Journal,36*, 1, 103-115.

National Task Force on Educational Technology (1986). Transforming American education: Reducing the risk to the nation. *Tech Trends, 31*,4, 12-24,35.

Romiszowski, A.J. (1981). *Designing Instructional Systems*. New York. Nichols Publishing.

Salomon, Gavriel. (1970). What does it do to Johnny? A cognitive-functionalistic view of research on media. *Viewpoints,46*,5.

Salomon, G., & Clark, R.E. (1977). Reexamining the methodology of research on media and technology in education. *Review of Educational Research, 47,* 99-120.

Simonson, Michael R. (1987). *"...Mere vehicles..." A symposium on the status of research in instructional technology.* Paper presented at the Annual Convention of the Association for Educational Communications and Technology, Atlanta, GA. ERIC Document Reproduction Service No.ED 285519.

Tennyson, Robert D. (1987). *Computer-based enhancements for the improvement of learning.* Paper presented at the Annual Convention of the Association for Educational Communications and Technology, Atlanta, GA. ERIC Document Reproduction Service No. ED 285522.

Winn, William. (1987). *Are media merely vehicles for instruction?* Paper presented at the Annual Convention of the Association for Educational Communications and Technology, Atlanta, GA. ERIC Document Reproduction Service No.ED 285523.

Selected and Reviewed Studies from Meta-Analysis Critique

Axeen, M.E. (1971). *Teaching library use to undergraduates--comparison of computer-based instruction and the conventional lecture.* Urbana, IL:University of Illinois. ERIC Document Reproduction Service No.ED 014 316.

Broh, C.A. (1975). *Achievement and attitude with computer related instruction: A field experiment.* Paper presented at the annual meeting of the American Political Science Association, San Francisco. ERIC Document Reproduction Service No. ED 110 399.

Cox, S. (1974). Computer assisted instruction and student performance in macroeconomic principles. *The Journal of Economic Education, 6* (1), 29-37.

Chiang, A., Stauffer, C., & Cannara, A. (1978). *Demonstration of the use of computer-assisted instruction with handicapped children. Final Report.* (Report No. 446-AH-60076A), Arlington, VA:RMC Research Corp. ERIC Document Reproduction Service No. ED 166 913.

Coffman, W.E. and Olsen, S.A. (1980). *The First Two Years of PLAN: An Evaluation of Program Impact.* Iowa City, IA:Iowa Testing Programs. ERIC Document Reproduction Service No. ED190 674.

Green, C., & Mink, W. (1973). *Evaluation of computer simulation of experiments in teaching scientific methodology.* St. Paul, MN: Macalester College. ERIC Document Reproduction Service No. ED 082 475.

Jones, L.A., & Sorlie, W.E. (1976). Increasing medical student performance with an interactive, computer assisted appraisal system. *Journal of Computer Based Instruction,2,*(3), 57-62.

McEwen,N., & Robinson, A. (1976). *Computer assisted instruction in secondary school French. Final Report.* Edmonton: Alberta University. ERIC Document Reproduction Service No. ED 150 846.

Paden, D.W., Dalgaard,B.R., & Barr, M.D. (1977). A decade of computer-assisted instruction. *The Journal of Economic Education, 9*(1), 14-20.

Skavaril, R.V. (1974). Computer based instruction of introductory statistics. *Journal of Computer Based Instruction, 1*(1),32-40.

Skavaril, R.V., Birky, J. Jr., Dahrkopf, R.E., & Knight, J.A. (1976). The use of CAI to provide problems for students in introductory genetics. *Journal of Computer Based Instruction, 3*(1),13-20.

Suppes, P., & Morningstar, M. Computer-assisted instruction. *Science, 166*, 343-350.

Title:

Learners Self-reports of Characteristics Related to Academic Achievement and Motivation

Author:

Ann R. Igoe

Learners Self-reports of Characteristics Related to Academic Achievement and Motivation

The relationship between academic achievement and motivation has been well established. Continuing motivation for a task has been defined as the future pursuit of that task under free choice conditions (Maehr, 1976). A student's beliefs about attribution, approval, challenge, self-competence, and locus of control have also been shown to have an impact on academic achievement and motivation.

Continuing motivation or return to task in school settings has been examined across a number of variables, including academic achievement, task difficulty, evaluation mode and gender. It has been theorized that challenge or mastery is a greater factor in boys' motivation to return to task, thus causing them to choose difficult tasks more frequently than girls, and that social approval is a greater factor in girls' motivation for a task, causing them to choose easier tasks more frequently (Harter, 1978; Van Hecke, Tracy, Cotler & Ribordy, 1984). While this adult reasoning seems logical, no self-report evidence has been presented from boys and girls themselves to validate it.

Academic achievement can be a major source of approval and social acceptance for females (Diesterhaft & Gerken, 1983; Gould & Slone, 1982; Prawat, 1976). Girls may be as desirous of challenge as boys, but they may sacrifice challenge for approval when the two conflict (Story & Sullivan, 1986; Van Hecke, et al., 1984; Whitehead, Anderson & Mitchell, 1987).

Other factors besides challenge and approval are often cited as being related to motivation. They include perceived competence, attribution of success and failure, and locus of control for learning activities. These factors have also been examined across a number of variables including gender and academic achievement.

Self-perception of competence is commonly accepted as a predictor of academic achievement. Females often show lower expectations of task success and lower ability estimates than do males (Jagacinski & Nicholls, 1987; McCombs, 1984; Prawat, 1976; Sleeper & Nigro, 1987). McCombs (1984) reports the growing recognition of the importance of students self-perceptions in the choices of learning activities which they pursue. Females attribute failure to lack of ability more often than do males (Miller, 1986; Parsons, Meece, Adler & Kaczala, 1982). Locus of control for responsibility of learning or task outcome is seen as an important correlate to self concept and to attributions for success (McCombs, 1984; Prawat, 1976; Uguroglu & Walberg, 1986).

Each of the individual factors of challenge, approval, competence, attribution and control have been related to one or more of the other factors in previous studies of continuing motivation (Bogie & Buckhalt, 1987; Diesterhaft & Gerken, 1983; Harter, 1981; Rosen & Aneshensel, 1978). However, no one to date has systematically collected descriptive data by sex and grade level on these five factors.

The purpose of this study was to investigate continuing motivation for hard and easy tasks and subjects' judgments of selected learner characteristics purported to be related to learning and motivation. The data were collected through subjects' self-reports on a Student School and Work Survey questionnaire, rather than being inferred from subject's behavior patterns in experimental research. A sample of approximately 100 males and 100 females from the seventh, ninth and eleventh grades was used to provide a relatively stable analysis of patterns by sex and grade level.

Method

Subjects

The subjects for this study were a total of 632 students, 338 females and 294 males, from grades seven, nine, and eleven at a junior and senior high school in a large suburban school district in the southwest. The numbers of subjects by grade level were 106 females and 93 males from the seventh grade, 128 females and 96 males from the ninth grade, and 104 females and 105 males from the eleventh grade. All subjects were from required classes. The junior high school was a feeder school for the high school. The student population of both schools ranges from lower to upper-middle socioeconomic class.

Procedures

This study was part of a larger survey which was developed for this study. It was administered by doctoral students to intact classes on a single day in each school. Students completing the survey were enrolled in required junior high math classes and required high school history classes. The survey was administered during the first twenty minutes of each class period.

Criterion Instrument

Most experimental studies of return to task (continuing motivation) have used relatively simple non-academic tasks such as anagrams, word search puzzles, and create-a-word problems. Tasks for this study were designed to be somewhat more substantial and applied. One involved helping to produce the school newspaper, with the implication that it involves writing, editing and page layout. The other involved coming up with an invention or creative idea.

The return to task measure consisted of scenarios in which a male or female character performed a task that the character considered to be either hard or easy. Eight scenarios were developed: 2 (easy or hard task difficulty) x 2 (male or female character) x 2 (newspaper or creative idea tasks). The presentations were counter-balanced for sex of character, task difficulty and type of task, and were randomly ordered by subject to prevent possible order bias.

Each subject responded to two scenarios, one with a female character and one with a male character. The number of scenarios per subject was limited to two in order to minimize the possibility that subjects would detect a pattern of change in the scenarios and respond more to the pattern than to the situation portrayed in each scenario per se. Across his or her two tasks each subject read about one boy and one girl, one school newspaper task and one creative task, and one easy and one hard task.

For each scenario, a subject answered one question about whether the scenario character would return to the task when offered another opportunity and a second question about whether the subject (i.e., the respondent) him/herself would return to the task in the same situation. 'Yes' responses were scored one point and 'no' responses were scored zero.

The Learner Characteristics Measure (Part Two) contains 15 questions. School related aspects of each of the five learner characteristics of approval, attribution, challenge, competence, and locus of control were assessed by three individual questions on 4-point Likert scales. Reliability data for each three-item subscale are reported in the results section.

Data Analysis

The return to task data were analyzed by a 2 x 2 x 2 x 3 (task difficulty, target individual, sex of subject and grade level) repeated measures MANOVA. Sex and grade level were between-subjects factors and task difficulty and target individual were within-subjects factors.

For the Learner Characteristics measure means were computed for each individual question and for the five subscales of approval, attribution, challenge, competence and locus of control. ANOVAs were performed for the five subscales and Scheffe post hoc tests were performed as appropriate. Cronbach's coefficient alpha was computed as a reliability measure for each of the subscales and Pearson product moment correlation coefficients were computed for each of the fifteen questions.

Results

Return to Task

Table 1 reports the return-to-task rate by task difficulty, target individual (scenario character or subject), sex of subject and grade level. The variable reflecting the greatest difference was task difficulty, on which the overall return rate was .85 for easy tasks and .42 for hard ones. Return rate by sex of subject was .66 for females and .61 for males. Return rates for the target individual were identical, .64 for both the scenario character and the respondent. Return rates were also very similar by grade level, .65 for grade 7, .62 for grade 9, and .64 for grade 11.

The 2 x 2 x 2 x 3 MANOVA revealed two significant differences for main-effect variables. The mean return rate of .85 for easy tasks was significantly higher than the mean of .42 for hard tasks, F (1,626) = 492.63, p < .0001. The mean rate of .66 for females was significantly higher than the mean of .61 for males, F (1,626) = 6.85, p = .009. Mean return rates did not differ significantly for target individual or for grade level.

The MANOVA yielded a significant two-way interaction for task difficulty by target individual, F (1,626) = 106.60, p < .0001. This interaction reflects the fact that the projected return to easy tasks was higher for the scenario character (.92) than for the subject (.79), whereas the projected return to hard tasks was higher for the subject (.49) than for the scenario character (.35). The task difficulty by target individual interaction is diagrammed in Figure 1.

The MANOVA also yielded significant two-way interactions for sex of subject by target individual and sex of subject by task difficulty. The sex of subject by target individual interaction, F (1,626) = 4.21, p = .041, shows females (.64) and males (.63) projecting very similar return rates for the scenario character, but females (.68) reporting a higher self-return rate than males (.59). The sex of subject by task difficulty interaction, F (1,626) = 6.92, p = .009, reflects a higher return rate to easy tasks for females (.91) than for males (.80), but identical rates for females (.42) and males (.42) to hard tasks.

Learner Characteristics

Part Two of the questionnaire dealt with the learner characteristics of approval of others, attribution of success or failure, desire for challenge, perception of self competence and locus of control in a learning situation. Table 2 shows the overall mean total for each characteristic based on scoring of zero to three. Zero was assigned as the score for the least positive response and three for the most positive response. Table 3 summarizes the statistically significant grade by sex ANOVAs for each of the five learner characteristics.

APPROVAL For approval, the overall mean was 2.24 for females and 1.94 for males. The higher score for females indicates more desire for the approval of others, and the lower score for males reflects less desire for approval. The sex-by-grade analysis of variance revealed that the difference between females and males was statistically significant, F (1,620) = 37.827, p < .001.

The overall means for approval by grade levels were 2.21 for grade seven, 1.94 for grade nine, and 2.02 for grade eleven, indicating that approval became less important across grade level. The ANOVA for grade level was statistically significant, F (2,620) = 5.715, p < .01. A Scheffe post hoc test revealed a significant difference between the seventh and eleventh grades F (2,629) = 5.659, p < .01, but not between any other grades. The interaction for sex by grade level was not statistically significant for approval nor for any of the other learner characteristics discussed hereafter.

ATTRIBUTION For attribution, the overall mean was 2.00 for females and 2.04 for males, reflecting higher attribution to themselves than to their teachers for success in school. The difference in means between females and males for attribution was not statistically significant.

The overall means for attribution across grade levels were 2.17 for grade seven, 2.03 for grade nine, and 1.87 for grade eleven, showing a trend toward less personal attribution as grade level increased. The difference by grade was statistically significant, F (2,619) = 23.720, p < .001. Scheffe post hoc results revealed that each grade was significantly different from the others.

CHALLENGE For challenge, the overall mean was 1.67 for females and 1.53 for males, indicating a significantly higher desire for challenge by females than by males, F (1,620) = 8.367, $p < .01$.

The overall means for challenge by grade level were 1.72 for grade 7, 1.60 for grade 9, and 1.48 for grade 11, showing less desire for challenge as grade level increased. The ANOVA revealed a significant difference by grade level, F (2,619) = 8.241, $p < .001$. The Scheffe test revealed that seventh grade was significantly different than the eleventh grade at the .001 level and the other differences were not significant.

COMPETENCE For competence, the overall mean was 1.99 for females and 2.01 for males, a non significant difference by sex indicating relatively high levels of self-competence for school tasks. The overall means by grade level for competence were 2.12 for grade 7, 2.10 for grade 9, and 1.79 for grade 11 showing lower feelings of self-competence as grade level increased.

The difference by grade level was statistically significant, F (2,619) = 28.988, $p < .001$, with the Scheffe test revealing that the eleventh grade was significantly lower than the seventh and ninth grades at the .001 level.

LOCUS OF CONTROL For locus of control, the overall mean was 1.94 for females and 1.98 for males, revealing a relatively high level of personal locus of control for both sexes. The overall means by grade level for locus of control were 1.97 for grade 7, 1.93 for grade 9, and 1.98 for grade 11. Neither the sex or grade level differences were statistically significant.

The reliability analyses for the five learner characteristics revealed alpha coefficients of .58 for approval, .29 for attribution, .47 for challenge, .61 for competence and .51 for locus of control. These coefficients are indicative of fairly high internal consistency for the three-item scales for all of the characteristics except attribution. The less acceptable value of .29 for attribution was a function of the negative wording of attribution item 2, "When you have trouble understanding something that the teacher is explaining, whose fault is it?". Clearly this item should be revised to be more consistent in wording with the other two attribution items prior to any further use of the scale.

Figure 2, shows the response patterns in graphic form for each of the five learner characteristics by grade level and sex. It can be seen that the characteristics show rather tight patterns of consistency of score levels, especially by sex and to a somewhat lesser degree by grade level.

Discussion

One purpose of this study was to investigate the return-to-task rates on hard and easy tasks projected by the subjects for scenario characters and reported for themselves personally. The data revealed a significantly higher rate of return to easy tasks over hard tasks and a significantly higher rate of return for females than for males. Three significant two-way interactions occurred. A higher return rate for easy tasks was reported for the scenario character and a higher return rate for hard tasks was reported for the subject. Females had significantly higher reported return-to-task rates for self-return and return to easy tasks than males.

The return rate for easy tasks (.85) was double the rate for difficult tasks (.42). These data are consistent with the common research finding that subjects prefer easy tasks over hard ones. However, the differences between the return rates for easy and hard tasks in this study were much greater than those reported in most other studies (e.g., Hughes, Sullivan, & Beaird ,1986 and Story & Sullivan, 1986).

The higher overall return-to-task rate for females than for males (.66 to .61) is a function of females' higher return rates for easy tasks, as reflected in the sex of subject by task difficulty interaction. The female return rate was significantly higher than the male rate (.90 to .80) on easy tasks but the same on hard tasks (.40 for both sexes). Thus, the higher overall rate for females was accounted for entirely by the difference on easy tasks. Most of this difference on easy tasks occurred on self-return (.87 for females and .70 for males), thereby yielding the sex of subject by target individual interaction in which females had a significantly higher self-return rate than males (.68 to .59) but projected nearly the same character return rate as males (.64 to .63).

This higher female return rate for easy tasks is consistent with prior research, but the similar return rates for males and females for difficult tasks are not. Higher return rates for females to difficult tasks have often been attributed to a greater desire for social approval by females and the probability that they will perform better on easier tasks (Harter, 1978; Story & Sullivan, 1986; Van Hecke, et al., 1984). Males have generally been found to be more willing to return to difficult tasks than females. The common interpretation of this finding is that males desire challenge more than females do.

Most of the research which shows a higher return rate for males has involved forced-choice situations in which subjects must choose between a hard and easy task. The present research raises the possibility that females may select the easy task more often because they have a greater preference for easy tasks than males do. But when given a choice of either doing a hard task or not doing it (rather than choosing between a hard or an easy task), females chose hard tasks for the scenario character and for themselves at about the same rate as males. Thus, the evidence from the present research does not indicate a greater preference for challenge by males than by females under the conditions of this study.

Two factors may help to explain the greater difference in return rates between difficult and easy tasks in this study than in others. One is that the tasks in this study were explicitly identified as being difficult or easy in the eyes of the scenario character. In contrast, in most studies of difficulty level, task difficulty is treated as an implicit variable that students must infer if they attend to it at all. Subjects may not always have personalized the scenario characters' perception of the difficulty of the task. Further, students in most studies report return to task only for themselves and not for another individual too. For example, some subjects may have reasoned, "Well, he thought that coming up with a creative idea was hard, but it wouldn't be hard for me." In addition, some subjects may have been attracted to a task itself, which could cause them to return to it irrespective of the character's judgment of its difficulty. A second factor is that the difference for subjects themselves (.79 for hard tasks and .49 for easy ones) was much less than the difference for scenario characters (.92 for hard and .35 for easy), though still higher than the reported rates from most continuing motivation research.

The results were highly consistent across grade levels. Subjects in all three grades showed the same high preferences for easy tasks and relatively low preferences for hard tasks, as well as the differential patterns by both target individual and sex. These data suggest that return to task preferences related to task difficulty and sex remain quite stable across the grade levels (7 to 11) and mean ages (12 years 9 months through 16 years 9 months) represented in this study.

On the Learner Characteristics measure, the high mean overall scores by sex of 2.24 for females and 1.94 for males for approval of others about school performance reveal a desire for approval. Approval of others was important for all subjects, but significantly more so for females at all grade levels. These results complement those of other researchers (Diesterhaft & Gerken, 1983; Prawat, 1976), who have found that academic achievement is a major source of approval and social acceptance for females.

The desire for approval decreased from the seventh grade to the ninth and eleventh grades. A similar pattern was detected by Harter (1975) with elementary school children and by Prawat (1976) with junior high school children. Apparently, a more internal orientation reflecting less desire for the approval of others emerges as children grow older and become more self-sufficient.

The overall results for attribution in school settings show high personal attributions for success for both sexes across grade levels, a result which agrees with other researchers (Parsons et al., 1982; Whitehead et al., 1987; Whitley & Frieze, 1985). The attribution results also revealed significantly weaker personal attributions as grade level increased. This finding differs from the meta-analysis results of Whitley & Frieze (1985), who found no significant grade level correlations, and contrasts with data obtained by Prawat (1976), who found that personal attributions increased significantly from sixth grade to seventh grade and again from seventh grade to eighth grade.

Prawat (1976) and Whitley and Frieze (1985) used different types of tasks from the school-related tasks in the present study. Prawat's (1976) tasks were a battery of affective tests not specifically related to school tasks. Academic tasks made up only 12% of the studies reported by Whitley and Frieze (1985). A possible reason for the different results in the present study is that children may have different attribution patterns in school than in other situations. In fact, Whitley and Frieze (1985) recommended research on naturalistic tasks in educational settings as an extension of their results.

The results for challenge revealed a significantly greater preference by females for challenge in the school-related tasks used in the survey. Current literature generally reports that males prefer more challenge than females. Most of the studies which have shown males desiring more challenge than females have used word search puzzles, create-a-word tasks, and anagrams, all of which could be considered by students as more like games than learning tasks (Bogie & Buckhalt, 1987; Harter, 1978; Miller, 1986). The school-related nature of the challenge questions in the present study may have appealed more to females than to males and resulted in their higher preference for challenge.

The desire for challenge decreased significantly as the grade level increased for both females and males. No explanations for this decrease can be found in current literature. Perhaps this decrease in desire for challenge reflects a greater occupation with nonschool matters. Children may become more bored with school tasks or more oriented toward social activities and work opportunities as they become more independent.

Both females and males showed high levels of self-competence for school-related tasks. The comparable levels across sex are contrary to some research findings which show lower ability estimates for females than males (Jagacinski & Nicholls, 1987, McCombs, 1984; Sleeper & Nigro, 1987), but consistent with others (Bogie & Buckhalt, 1987; Prawat, 1976).

Very few studies have addressed grade level and self-competence. In one such study, Prawat (1976) found no significant differences for grade as a source of variance for a self-esteem measure. In the present study, competence tends to decrease as grade level increases.

The results for locus of control show high internal ratings that remained near the same level across grade levels for both females and males, a result which is consistent with several studies reported by Stipek and Weisz (1984). The present study measured the subjects' desired locus of control for school activities, either self-determined or teacher-determined. Most researchers have assessed locus of control and its relationship to self-competence and achievement using more general measures. They have typically found a positive correlation between internal locus of control and high self-competence (Diesterhaft & Gerken, 1983; Prawat, 1976; see also Stipek & Weisz, 1981 for a review).

A summary of the results by sex across the five learner characteristics indicates that females showed a higher desire than males for approval and challenge in school-related tasks. The result for approval is consistent with other research findings, but the result for challenge contrasts with most research data. The most plausible explanation for the higher female preference for challenge appears to be the emphasis on school-type tasks in the present survey as contrasted with the more puzzle-like tasks in other studies.

The results for learner characteristics across grade levels reveal a pattern in which older students take less responsibility for their learning, desire less personal challenge, and feel less school-related competence. The opposite trend for school-related approval, which becomes more internal as grade level increases, fits this pattern as well. Other researchers have reported similar findings for achievement measures (Prawat, 1976) and for an intrinsic/extrinsic classroom orientation measure (Harter, 1981). Prawat (1976) suggests the increasing importance of peer groups as a possible explanation of his results. Contrarily the approval of friends received a very low rating in the present study. As students grow older, they become more independent and often look more to the world outside school. It may be that school, and academic activities, do not have as much importance in students' lives as they did when the students were younger (Harter, 1981).

Two overall patterns for females appeared across both parts of the survey. Approval of others is important for females as evidenced by their desire to return to easy tasks and their high approval scores on the learner characteristics scale, However, females also showed a stronger desire for difficult tasks and challenge in school activities than found by other researchers (Harter, 1978; Miller, 1986).

Boys and girls had identical self-reported return rates to tasks that the scenario characters thought were difficult -- tasks that therefore could be considered as potentially challenging. Conscious efforts by educators to provide greater social approval to girls for selecting challenging tasks could possibly increase their frequency of selection of such tasks. Research which investigates the effects of providing greater social approval for selecting and participating in challenging tasks could lead to better understanding of the patterns of continuing motivation of boys and girls for various types of school tasks.

References

Bogie, C. E. & Buckhalt, J. A. (1987). Reactions to failure and success among gifted, average, and EMR students. Gifted Child Quarterly, 31(2), 70-74.

Diesterhaft, K. and Gerken, K. (1983). Self-concept and locus of control as related to achievement of junior high students. Journal of Psychoeducational Assessment, 1, 367-375.

Gould, R. J. and Slone, C. G. (1982). The "Feminine Modesty" effect. A Self-presentational interpretation of sex differences in causal attribution. Personality and Social Psychology Bulletin, 8, 477-485.

Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. Developmental Psychology, 17, 300-312.

Harter, S. (1978). Pleasure derived from challenge and the effects of receiving grades on children's difficulty level choices. Child Development, 49, 788-799.

Harter, S. (1975). Mastery motivation and need for approval in older children and their relationship to social-desirability response tendencies. Developmental Psychology, 11, 186-196.

Hughes, B., Sullivan, H. J., and Beaird, J. (1986). Continuing motivation of boys and girls under differing evaluation conditions and achievement levels. American Educational Research Journal, 23, 660-667.

Jagacinski, C. M. and Nicholls, J. G. (1987). Competence and affect in task involvement and ego involvement: The impact of social comparison information. Journal of Educational Psychology, 79, 107-114.

Maehr, M. L. (1976). Continuing motivation: An analysis of a seldom considered educational outcome. Review of Educational Research, 46, 443-462.

McCombs, B. L. (1984). Processes and skills underlying continuing intrinsic motivation to learn: Toward a definition of motivational skills training interventions. Educational Psychologist, 19, 199-218.

Miller, A. (1986). Performance impairment after failure: Mechanism and sex differences. Journal of Educational Psychology, 78, 486-491.

Parsons, J. E., Meece, J. L., Adler, T. F., & Kaczala, C. M. (1982). Sex differences in attributions and learned helplessness. Sex Roles, 8, 421-432.

Prawat, R. S. (1976). Mapping the affective domain in young adolescents. Journal of Educational Psychology, 68, 566-572.

Rosen, B. C. & Aneshensel, C. S. (1978). Sex differences in the educational-occupational expectation process. Social Forces, 57 (1), 164-186.

Sleeper, L. A. and Nigro, G. N. (1987). It's not who you are but who you're with: Self-confidence in achievement settings. Sex Roles, 16, 57-69.

Stipek, D. J. and Tannatt, L. M. (1984). Children's judgments of their own and their peers' academic competence. Journal of Educational Psychology, 76, 75-84.

Story, N. O. and Sullivan, H. J. (1986). Factors that influence continuing motivation. Journal of Educational Research, 80, 86-92.

Uguroglu, M. E. & Walberg, H. J. (1986). Predicting achievement and motivation. Journal of Research and Development in Education, 19 (3), 1-12.

Van Hecke, M., Tracy, R. J., Cotler, S., and Ribordy, S. C. (1984). Approval versus achievement motives in seventh-grade girls. Sex Roles, 11, 33-41.

Whitehead, III, G. I, Anderson, W. F., and Mitchell, K. D. (1987). Children's causal attributions to self and other as a function of outcome and task. Journal of Educational Psychology, 79, 192-194.

Whitley, B. E., Jr. & Frieze, I. H. (1985). Children's causal attributions for success and failure in achievement settings: A meta-analysis. Journal of Educational Psychology, 77, 608-616.

Table 1

Return to Task Rate by Task Difficulty, Target Individual, Sex of Subject, and Grade Level

	Grade							
	7		9		11		Total	
Difficulty Level	F	M	F	M	F	M	F	M
Character Return								
Easy	.93	.91	.91	.90	.98	.88	.94	.90
Hard	.35	.37	.35	.30	.34	.38	.35	.35
Total	.64	.64	.63	.60	.66	.63	.64	.63
Self Return								
Easy	.85	.74	.87	.71	.88	.66	.87	.70
Hard	.54	.49	.47	.45	.46	.52	.49	.49
Total	.70	.61	.67	.58	.67	.59	.68	.59

Overall Mean Return Rates

Hard = .42
Easy = .85
Male = .61
Female = .66
Character = .64
Self = .64
Grade 7 = .65
Grade 9 = .62
Grade 11 = .64

Figure 1. Return to Task Rates for Task Difficulty by Type of Character

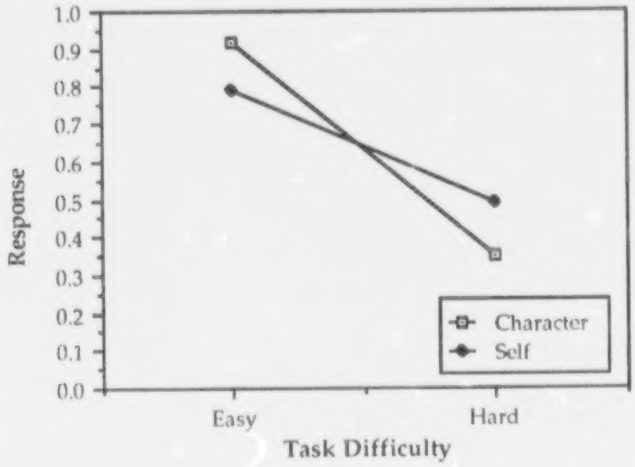


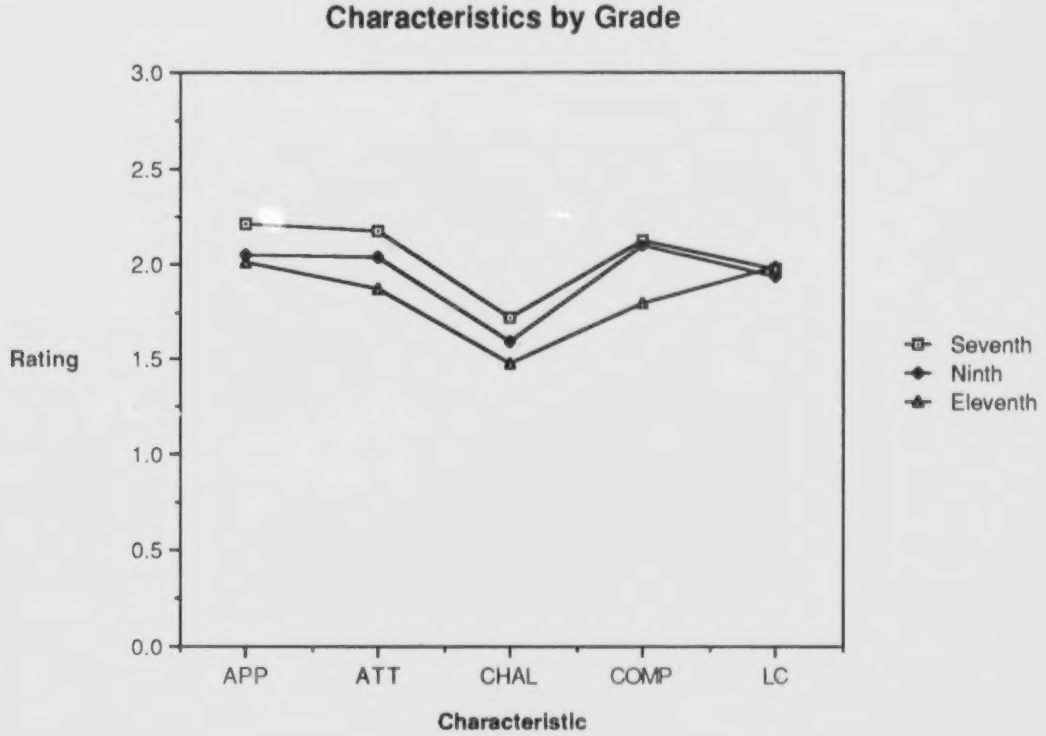
Table 2 Overall Means for Each Learner Characteristic

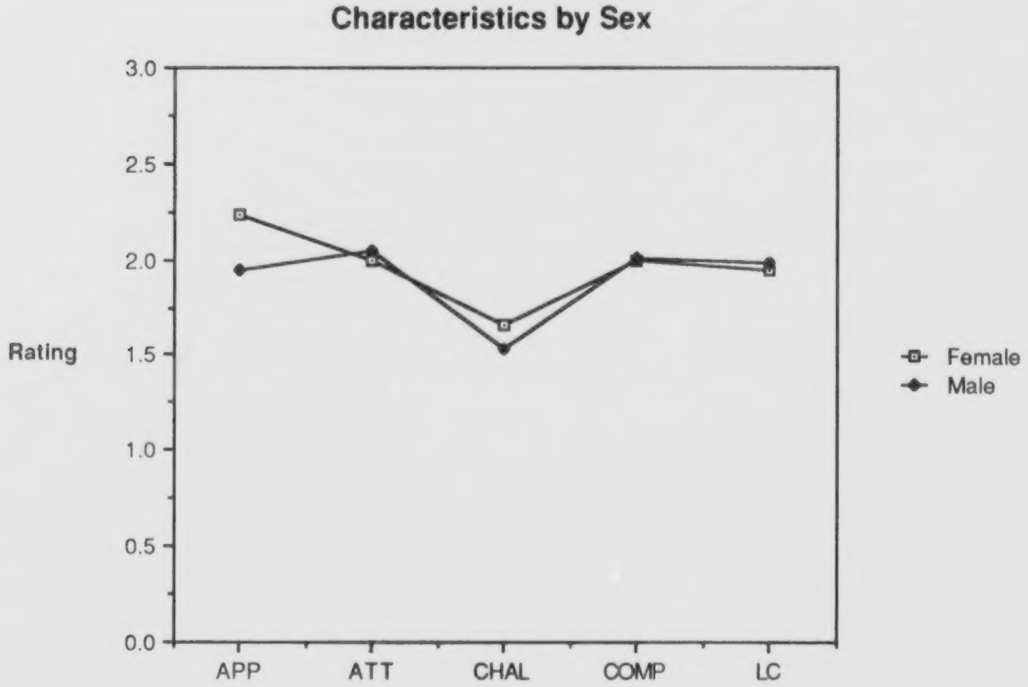
	Grade							
	7		9		11		Total	
Characteristic	Female	Male	Female	Male	Female	Male	Female	Male
Approval	2.39	2.03	2.20	1.89	2.13	1.90	2.24	1.94
Attribution	2.15	2.19	2.02	2.04	1.84	1.90	2.00	2.04
Challenge	1.80	1.64	1.68	1.51	1.51	1.45	1.67	1.53
Competence	2.17	2.07	2.07	2.12	1.73	1.85	1.99	2.01
Locus Control	1.95	1.99	1.93	1.93	1.94	2.02	1.94	1.98

Table #3 Summary of Significant Differences for Learner Characteristics

Characteristic	Significance	Variable	Difference
Approval	.001	Female	Females rated approval more important
	.01	Grade	7th graders rated approval more important
Attribution	.001	Grade	7th graders rated internal attribution more highly
Challenge	.01	Female	Females rated challenge higher
	.001	Grade	7th graders rated challenge higher
Competence	.001	Grade	7th graders rated competence higher than 11th graders
Locus of Control			There were no significant differences

Figure 2. Learner Characteristcs by Grade and Sex.





Title:

The Effects of Pictorial Complexity and Cognitive Style on Visual Recall Memory

Authors:

Romaine R. Jesky
Louis H. Berry

THE EFFECTS OF PICTORIAL COMPLEXITY AND COGNITIVE STYLE ON VISUAL RECALL MEMORY

Romaine R. Jesky
Geneva College
Beaver Falls, PA

Louis H. Berry
University of Pittsburgh
Pittsburgh, PA

Research in the use of visuals for instructional purposes has increasingly identified the interaction between the learner's individual cognitive skills and the design factors incorporated in the instructional method. Researchers investigating memory processes have theorized multiple processing systems for different modalities of information, with the visual, verbal and auditory being of particular interest to learning theorists (Levie & Levie, 1975; Paivio, 1978; Winn, 1980, 1982). Predictably, research interest has led to detailed explorations of those factors within the area of visual learning which contribute to the effective processing of pictorial material.

Considerable research has addressed the factor of visual complexity in instruction (Dwyer 1972, 1978, 1987). This research was drawn from a larger theoretical controversy which continues regarding visual complexity and human information processing. It has long been contended that the mere addition of visual cues will increase the ability of the viewer to store and retrieve visual information.. This orientation, termed "realism theory" by Dwyer (1967), has strong theoretical antecedents (Dale, 1946; Morris, 1946; Carpenter, 1953 and Gibson, 1954) and is indeed the major premise of cue summation theory (Severin, 1967). Other researchers (Broadbent, 1958, 1965; and Travers, 1964) however, took strong opposition to this theoretic base on the grounds that the human information processing system is of limited capacity and consequently, in times of rapid information reception, irrelevant cues may block the processing of other, more relevant information. Studies (Kanner, 1968; Katzman and Nyenhuis, 1972; Dwyer, 1972, 1978, 1987) have investigated this apparent contradiction with conflicting results.

During the past thirty years, the concept of cognitive style has emerged as an area of substantial research. Messick (1976) defined cognitive style as "consistent patterns of organizing and processing information". The specific cognitive style of field dependence, identified by Witkin, Oltman, Raskin and Karp (1971) is generally defined as the differential ability of individuals to overcome figural embeddedness. This perceptual ability is considered to be representative of a more global ability to impose structure upon perceived information.

A number of researchers have investigated how the aptitude of field dependence relates to an individual's ability to process pictorial information. French (1983) found that field independent subjects experienced less difficulty processing unusually complex material than did field dependent viewers. Wieckowski (1980) found a significant difference in favor of field independent subjects in a pictorial recognition task involving color. These findings, relative to a recognition task, were further substantiated in research conducted by Berry (1984).

More limited research has addressed the effects of visual complexity on recall memory. Sampson (1970) found pictures better than words in both immediate and delayed free recall situations. Ritchey (1982) reported an advantage in recall for outline drawings over detailed drawings. Conversely, Alfahad (1990) found realistic color visuals to be superior to black and white or line drawing visuals in a recall memory task. Recall memory involving the interaction of pictorial complexity and field dependence has, however, not been investigated.

The purpose of this study was to explore the effect of the interaction between cognitive style differences (field dependence/field independence) and various degrees of visual complexity on pictorial recall memory.

METHOD

The stimulus materials used in this study were three sets of visuals, each produced in three visual formats; line drawing, black & white and color. To create the sets of visuals, three different collections of common household items (32 per set) were randomly arranged on a neutral photo backdrop. In selecting the objects, care was taken to ensure that no verbal labels, names or symbols were visible. Each set was photographed on color slides and then later recopied onto black and white slides. A line drawing of each scene was created by an artist working from the projected black and white slides. These drawings were also copied onto 35mm slides. The resulting stimulus materials consisted of three sets of slides, each composed of three different treatment versions of the same image; one rendered in photographic color, another in a black and white photographic format and the third in a line drawing format. The materials were validated by comparison with the Visual Memory Test developed by Salomon & Cohen (1977).

This test was intended to measure an individual's ability to recall a number of objects from a visual stimulus which was presented in a line drawing format. The reported reliability of the test was .74. Pearson Product Moment Correlation Coefficients were calculated between the Visual Memory Test and the three treatments developed for this study. The obtained r values were line drawing-.464, black and white-.469, and color-.504. Each of these values was determined to be significant at the .01 level. Further validation was achieved via regression analysis where the Visual Memory Test was found to explain 21.5% of the variance of the line drawing treatment, 25.4% of the variance of the black and white treatment and 21.9% of the variance of the color treatment.

For purposes of presentation, the slides were organized into sets of three different treatments which could be presented to three separate experimental groups. The ordering of the sets of slides for the study is shown in Table 1. By rotating the slides in this manner, each different set of slides was counterbalanced with each treatment resulting in a repeated measures design which provides greater experimental precision.

Table 1
Ordering of Stimulus Materials

Experimental Group	Color	Black & White	Line Drawing
Group 1 (Set 1)	Scene A	Scene B	Scene C
Group 2 (Set 2)	Scene C	Scene A	Scene B
Group 3 (Set 3)	Scene B	Scene C	Scene A

Subjects for the study were 86 undergraduate students from a liberal arts college enrolled in two core curriculum courses, thereby ensuring a wide representation of majors. The sample consisted of 56 females and 39 males. Any subjects with visual handicaps were eliminated from the sample.

In a class session prior to the experiment sessions, subjects were administered the Group Embedded Figures Test (Witkin et al., 1971). The mean score on this instrument was 10.17 with a standard deviation of 5.06. The Hampel estimate on this data was 10.44 which differed from the mean by less than 10% indicating that the GEFT scores of the subjects were normally distributed. Subjects were grouped into three levels by their GEFT scores (0-8, field dependent; 9-13, indeterminate; 14-18, field independent). The distribution of subjects is shown in Table 2.

Table 2.
Distribution of Subjects by Cognitive Style

Group	N
Field Dependent (1-8)	30
Indeterminate (9-13)	29
Field Independent (14-18)	27
Total	86

Based on the cognitive style scores, subjects were assigned to one of three groups by means of a stratified random sampling procedure. Each group was then randomly assigned to a particular set of stimulus slides. In a semi-darkened room, each group viewed an individual slide for a period of 20 seconds after which the slide was removed and the lights raised. Subjects then received four minutes to write down as many objects as they could recall from the slide. This procedure was repeated again with each of the other two slides assigned to the particular group of subjects with a rest period of five minutes provided between each slide A similar procedure was followed with each of the additional groups.

ANALYSIS

The design of the study was a 3 x 3 repeated measures type with three repeated levels of the visual complexity factor (color, black & white and line drawing) and three levels of the cognitive style factor (field dependent, indeterminate and field independent). Means and standard deviations for the three treatment groups by cognitive style levels is presented in Table 3.

Table 3.
Means and Standard Deviations for Visual Treatments by Cognitive Style Levels (N = 86)

Treatments	FD		IND.		FI		Marginal	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.
Color	10.16	2.01	10.41	1.99	10.96	2.00	10.50	2.00
Black & White	8.96	1.79	9.41	1.70	10.15	1.75	9.80	1.78
Line Drawing	8.83	1.93	8.52	1.81	9.55	1.79	8.95	1.84
Marginal	9.32	1.94	9.44	1.80	10.55	1.85	9.75	1.86

Analysis of variance procedures produced a significant main effect for visual complexity (F=12.20, p<.0001), but no significant interaction or main effect for cognitive style. *Post hoc* comparisons via the Scheffé procedure indicated significant differences between all three levels of the complexity factor in the following order: color > black & white > line drawing.

DISCUSSION

The findings indicated that the cognitive style factor of field dependence was not significantly related to recall memory under the varied levels of the visual factor. It was apparent, however, that recall memory was related to the visual presentation mode. The ordering of the means with color highest, followed by black & white and line drawing respectively, suggests that cue summation theory is a valid predictor of visual processing in recall memory. It is further suggested that the "realism" continuum identified by Dwyer is similarly a reliable descriptor of recall processing in that more realistic visuals facilitate recall memory better than do less realistic or visually complex materials.

Although recognition memory tasks are considered to demonstrate the most pure measure of the content of memory, recall tasks provide insight into the encoding and retrieval processes involved in memory (Loftus & Loftus, 1976). It has also been suggested that recall memory tasks involve two separate cognitive processes, a memory search and a recognition process (Anderson & Bower, 1972). This search process would depend on the cues employed in encoding the information in memory and consequently be strongly related to the types of visual cues incorporated in the stimulus. In the present study, it would appear that the color cues performed more effectively in encoding than did the monochrome cues . Similarly, the monochrome information was superior to the line drawing material. It can be concluded, then, that more realistic cues are more effective encoding devices than are less realistic cues. The reasons for this performance may rest in the fact that the realistic cues present a more complete representation of a schema already present in memory, making the encoding/storage process easier. Less realistic materials, on the other hand, may not provide as complete a representation and therefore may not encode as completely.

REFERENCES

Alfahad, F. N., (1990). The interactive effects of cognitive functioning and visual realism on visual memory recall tasks. (Unpublished doctoral dissertation, University of Pittsburgh).

Anderson, J. R. & Bower, G. H., (1972). Recognition and retrieval processes in free recall. *Psychological Review. 79.* 97-123.

Berry, L. H. (1984). *The role of cognitive Style in processing color information: A signal detection analysis.* Paper presented at the annual convention of the Association for Educational Communications and Technology, Dallas, TX.

Broadbent, D.E. (1958). *Perception and Communication.* New York: Pergamon Press.

Broadbent, D.E. (1965). Information processing in the nervous system.*Science. 35,* 457-462.

Carpenter, C.R. (1953). A theoretical orientation for instructional film research. *AV Communication Review, 1,* 38-52.

Dale, E. (1946). *Audio-visual methods in Teaching.* New York: Dryden Press.

Dwyer, F.M. (1967). Adapting visual illustrations for effective learning. *Harvard Educational Review,.37,* 250-263.

Dwyer, F.M. (1972). *A guide for improving visualized instruction..* State College, PA: Learning Services.

Dwyer, F.M. (1978). *Strategies for improving visual learning.* State College, PA: Learning Services.

Dwyer, F. M., (Ed.).(1987) *Enhancing Visualized Instruction - Recommendations for Practice* . State College: Learning Services.

French, M. (1983). *Learning with visuals through aptitude sensitive instruction.* Paper presented at the annual convention of the Association for Educational Communications and Technology, New Orleans, LA.

Gibson, J.J. (1954). A theory of pictorial perception. *A V Communications Review. 2,* 2-23.

Kanner, J.H. (1969). *The instructional effectiveness of color in television: A review of the evidence.* Palo Alto, CA: Stanford University.

Katzman, H. & Nyenhuis, J. (1970). Color vs black and white effects on learning, opinion and attention. *AV Communication Review. 20,* 16-28.

Loftus, G. R. & Loftus, E. F., (1976). *Human Memory: The processing of Information.* New York: Wiley.

Messick, S. (1976). *Individuality in Learning.* San Francisco: Jossey-Bass.

Morris, C.W. (1946). *Signs, language and behavior.* New York: Prentice-Hall.

Ritchey, G. H. (1982). Notes, comments and findings: Pictorial detail and recall in adults and children. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 2, 139-141.

Sampson, J. R., (1970). Free recall of verbal and non-verbal stimuli. *Quarterly Journal of Experimental Psychology.* 22. 215=221.

Severin, W. (1967). Another look at cue summation. *AV Communication Review.* *15,* 233-245.

Salomon, G. & Cohen, A. A., (1977). Television formats, mastery of mental skills,and the acquisition of knowledge. *Journal of Educational Psychology. 69.* 612-619.

Travers, R. M. G. (1964). The transmission of information to human receivers. *AV Communication Review. 12,* 373-385.

Wieckowski, T. J. (1980). *The interactive effects of color and cognitive style on a pictorial recognition memory task.* Paper presented at the annual convention of the Association for Educational Communications and Technology. Denver, CO.

Winn, W., (1980). Visual information processing: A pragmatic approach to the "imagery question". *Educational Communications and Technology Journal. 28.* 120-133.

Winn, W., (1982). Visualization in learning and instruction: A cognitive approach. *Educational Communications and Technology Journal. 30.* 3-25.

Witkin, H. A., Oltman, P., Raskin, E. & Karp, S. (1971). *A Manual for the Embedded Figures Test.* Palo Alto, CA: Consulting Psychologists Press.

Title:

Chaos, Communication, and Educational Technology

Author:

James W. King

Chaos, Communication, and Educational Technology

James W. King
University of Nebraska - Lincoln
Lincoln, NE 68583-0918

Introduction

"Chaos: Making of a New Science" was a best selling book. In it, Gleick (1987) offered the general public a glimpse of the scientific concept chaos. His illustrative overview proved very popular. Following closely came PBS programming from NOVA and Discover. Earlier, in 1984, Prigogine and Stengers wrote what was also to become a best selling book, "Order Out of Chaos." The Nobel Prize winner Prigogine integrated many scientific concepts with literature and nature to present "man's (sic) new dialogue with nature."

Why the interest in chaos? Surely the concern about scientific illiteracy is one reason. However, Gordon and Greenspan (1988) suggest another reason of more interest to communicators. "Chaos is the study of disorder that arises in nonlinear systems and should be of interest to technological, social and economic forecasters since the systems with which they deal are usually nonlinear and disorderly" (p. 1). The emerging framework (based on chaos and the work of Prigogine and Stengers) is intriguing and provocative. Toffler (1984) writes that "it shifts attention to those aspects of reality that characterize today's accelerated social change: disorder, instability, diversity, disequalibrium, nonlinear relationships (in which small inputs can trigger massive consequences) and temporality -- a heightened sensitivity to the flows of time" (p. xiv-xv).

Within this social change, communication systems evolve into complexity. The models based on those realities become more encompassing and possibly "oppressive" (Hayles, 1990). New means of dealing with complexity, combining perspectives, and asking "what if" questions become necessary.

Chaos gives communicators new data, suggests innovative approaches to old ideas, and reaffirms certain approaches. Both Jonassen (1990) and Winn (1990) recently discussed some implications of chaos for instructional design, and Jonassen described chaos as the "theory of complexity." Others have researched chaos and discussed war (Saperstein, 1984), information flow (Shaw, 1981, 1984), computational systems (Huberman, 1989), animal systems (Gould, 1986), computing (May, 1989), and technological and social forecasting (Gordon and Greenspan, 1988). Several papers use chaos to discuss education (Cziko, 1989; Jonassen, 1990; Lehrer, et al., 1990; Winn, 1990). For communication and educational communications and technology, chaos seems to offer some intriguing ways of seeing communication systems. Chaos presents us with new approaches to examining traditional problems, and promises new life to under-used techniques to study communication systems.

In this paper I will present a background on chaos, discuss concepts drawn from the wide-ranging literature of chaos, and list implications chaos could have for communicators. I will not discuss systems theory, an important concern for both communication and chaos. Salisbury (1990) has recently provided a good review of systems concepts for instructional development, and other general systems references are available (Ackoff, 1971; Ashmos and Huber, 1987; Boulding, 1956; Gharajedaghi and Ackoff, 1984). I will not discuss information or information theory. Dervin's theory of information as construction (1989) fits into the current cognitive perspective of communication and her work is excellent. Shaw (1981, 1984) furnishes an overview of information in dynamic systems. Hayles (1990) reviews Shannon's work (1949) and produces an interesting discussion on probability and information. I would recommend that curious readers pursue these references.

The Background: Chaos

Hayles (1990) suggests two ways to view chaos. The first is based on Prigogine and Stengers (1984) and focuses on order out of chaos. Problems develop and can be solved through the "spontaneous emergence of self-organization." The second, and perhaps more highly popularized view (Gleick, 1987), discusses order within chaos. Systems generate new information, the "deeply encoded structures" of strange attractors. While I will refer to the Prigogine and Stengers chaos, most of this paper will be of the "order within chaos" type. Implications will be drawn from both views. The following fourteen points are a brief background to understanding chaos.

1. Predictability relates to simplicity and regularity in systems theory, and, for some researchers, is a "measure of rationality" (Curtis, 1990). Unpredictability is associated with complexity and irregularity. During the past two decades, new ways to study complexity have emerged from many disciplines, such as physics, astronomy, and biology. These methods have grown into a field of study and methodology labeled chaos theory. Chaos "offers a way to understand complicated behaviors as something that is purposeful and structured instead of extrinsic and accidental" (Pool, 1989a, p. 245).

2. Mechanistic models suggest internal regularity, and cause and effect that is understandable and logical (Gharajedaghi and Ackoff, 1984). When cause and effect relationships are unclear, the involved components are described as random elements (Crutchfield, et al., 1986). Given enough data, researchers typically thought, predictability was possible. Now, it has been discovered that "simple deterministic systems with only a few elements can generate random behavior" (p. 46). This randomness, found to be fundamental to the system, is described as "chaos." Deterministic chaos, writes Mandelbrot (1990), shows that although simple systems may obey deterministic laws, their behavior is unpredictable.

3. "In principle, the future is completely determined by the past, but in practice small uncertainties are amplified, so that even though the behavior is predictable in the short term, it is unpredictable in the long term" (Crutchfield, et al., 1986, p. 46). Because initial systems conditions are often impossible to fully and accurately survey, chaos occurs (J. May, 1989).

4. This is due partly to quantum mechanics, which suggests that there is always a randomness (Heppenheimer, 1989) and uncertainty, especially in initial measurements (Crutchfield, et al., 1986). Chaos seems to multiply these uncertainties, resulting in a lack of good prediction ability. Because of the fundamental nature and scope of probability and randomness, those notions are thought to be a unifying principle (Chaitin, 1990). Chaos extends this principle universally.

5. Some of the background for chaos lies in dynamical systems. "A dynamical system is a system whose evolution from some initial state (which we know) can be

described by a set of rules" (Tsonis and Elsner, 1989). There are two parts of dynamical systems:

> 5.1. the essential data about the system, termed "state." This is often abstract data, which may be "...the degrees of freedom of the system...(Crutchfield, et al, 1986, p. 49). A state may be stable or unstable, and
> 5.2. rules describing how the state evolves with time, termed a "dynamic." Time may be either continuous time (flow) or discrete time (mapping). (Crutchfield, et al., 1986).

Stewart (1989) writes that "...chaos is a dynamic phenomenon. It occurs when the state of a system changes with time" (p. 24). Mathematics has four ways to characterize behavior as a function of time: steady states, oscillations, chaos, and noise (Glass and Mackey, 1988). However, Dunbar (1990) says that chaos "...indicates more a general category of behavior than a specific kind of movement in a dynamical system" (p. 4).

6. Timepaths generated by dynamical systems are very sensitive to changes in initial conditions. "Robust and reliable models are hard to develop" (Baumol and Benhabib, 1989). Timepaths may be characterized four ways:

> 6.1. oscillatory and stable
> 6.2. oscillatory and explosive
> 6.3. nonoscillatory and stable
> 6.4. nonoscillatory and explosive

7. Systems have attractors. An attractor "is what the behavior of a system settles down to, or is attracted to... A system can have several attractors. If that is the case, different initial conditions may evolve to different attractors" (Crutchfield, et al., 1986, p. 50). Thus, attractors take two general forms here: a) a single point that is like a steady state -- nothing happens, and b) a closed loop that is periodic, the behavior repeating itself again and again (Stewart, 1989).

Predictable attractors can be simple points. But attractors that repeat upon themselves are chaotic, or strange attractors. This repetition is *ad infinitum*, "a fractal, an object that reveals more detail as it is increasingly magnified" (Crutchfield, et al., 1986, p. 51).

8. Systems that are chaotic provoke an internal randomness, "...on their own, without the need for any external random input" (Crutchfield, et al., 1986, p. 53). This chaotic behavior "resembles randomness" (Savit, 1990, p. 49).

9. In predictable systems, future conditions or behaviors can be projected from beginning benchmarks. There is cause and effect. Such systems are not overly malleable to errors in the initial stages. However, chaotic attractors change or displace the initial data and put in new data. Small scale uncertainties become larger. "After a brief time interval, the uncertainly specified by the initial measurement covers the entire attractor and all predictive power is lost; there is simply no causal connection between past and future" (Crutchfield, et al., 1986, p. 53)

10. The "average rate" at which data is produced by a system determines its level of chaos. (Crutchfield, et al., 1986).

11. Chaos is grounded in mathematics (Pool, 1989a), particularly in non-linear formulations (May, 1976). The mathematics of chaos allows researchers to define

randomness triggered by simple deterministic dynamical systems (Tsonis and Elsner, 1989). Simply, chaos gives scientists, and now educators, the ability to see patterns in processes once considered totally random.

Glass and Mackey (1988) refer to noise as "chance fluctuations, a randomness that is completely unpredictable, except in a statistical sense" (p. 37). They differentiate noise from chaos, which is a randomness or irregularity in a deterministic system and can be found and identified in the complete absence of noise.

12. There are many definitions of chaos. In his book, Gleick (1987) provided over six definitions for chaos from the research literature. Each writer presents another definition. For communicators, Pool's (1986a) definition may be useful:

> "...deterministic randomness...a type of randomness that appears in certain physical and biological systems and is intrinsic to the system rather than caused by outside noise or interference" (Pool, 1989a, p. 25).

13. May (1976) discusses chaos from a systems perspective: "...the very simplest nonlinear difference equations can possess an extraordinarily rich spectrum of dynamical behavior, from stable points, through cascades of stable cycles, to a regime in which the behavior (although fully deterministic) is in many respects 'chaotic', or indistinguishable from the sample function of a random process" (p. 459).

14. Summarizing the literature of chaos, Gleick (1987) points out that simple systems give rise to complex behavior, and complex systems often can be explained by simple behaviors and rules.

Concepts in Chaos

Several principles and attributes of chaos are of interest to educational communications and technology. These include **sensitive dependence on initial conditions, nonlinearity, and complexity which includes fractals, bifurcations, and strange attractors.**

1. Sensitive dependence on initial conditions: Initial states are the beginning conditions in a system. The system has a "sensitive dependence" on these initial conditions. Pool (1989b,c) notes that if the initial configuration of a system is changed, no matter to what degree, its future behaviors can be altered radically. Uncertainty, the product of a quantum mechanics framework, means that initial conditions are always somewhat unstable or unknown. In chaotic environments, long-term systems solutions may be inherently unpredictable (Pool, 1989a; Pool, 1990b).

Dependence on initial conditions is not a new or alien idea. In communication situations, initial conditions include the problem or source description, audience identification and analysis, and first impressions in an encounter.

2. Nonlinearity: When the system's end product is proportionally equated to the input, over the total input range, the system is linear (Puccia and Levin, 1985). There is a correspondence between cause and effect. Proportionality results in straight lines. Small changes in causes lead to small changes in effects. In nonlinear systems, "the parameters of the system do not vary proportionally" (McRobie and Thompson, 1990).

Nonlinear systems may not proceed step by step and may not have simple dynamical properties. Nonlinear differential equations are used to describe the state. There are unanticipated events; erratic systems behavior is proven to have noting to do with outside input. Cause and effect are not corresponding. Small changes in initial conditions can lead to large changes in effects.

With nonlinearity, changes are not necessarily ordered or orderly. Prediction is questionable. Uncertainty is prevalent. Examples of nonlinearity would include the entire process of message design and delivery.

3. Systems that are chaotic have complex forms: This means those systems include **fractals, bifurcations, and strange attractors.**

3.1. Fractals: "Fractals are geometrical shapes that...are not regular at all. First, they are irregular all over. Second, they have the same degree of irregularity on all scales. A fractal object looks the same when examined from far away or nearby--it is self-similar. As you approach it, however, you find that small pieces of the whole, which seemed from a distance to be formless blobs, become well-defined objects whose shape is roughly that of the previously examined whole" (Mandelbrot, 1990, p. 39).

Systems' components may be self-similar and recursive. They may involve a scaling structure with patterns inside of patterns. While the details of the system are random, the fractal dimension is constant, no matter how complex or holistically viewed. That means that fractals have a "complete structure on any scale of magnification" (Stewart, 1989, p. 46). Fractals may be a "way of seeing infinity" (Gleick, 1987, p. 98). Hayles (1990) calls this approach a "shift in focus" (p. 13) from individual units to levels and systems perspectives. An example of fractals would be the difference in looking at details versus the patterns in systems.

3.2. Bifurcation point: A system undergoes bifurcation, a splitting and diverging of paths that cause system change, when the number and/or stability of a steady state and cycle change (Glass and Mackey, 1988). Bifurcations are real changes in qualitative dynamics, and can start a fractal process (Gordon and Greenspan, 1988), leading to chaos. In fact, Glass and Mackey (1988) write that "period-doubling bifurcations followed by irregular dynamics is taken as evidence that the irregular dynamics are chaotic" (p. 49).

Bifurcations in human systems can be seen in Flach's work (1988) on the life cycle and periods of stress. Stress on stable systems produces a disruption, a bifurcation, and the system changes, often unexpectedly and unpredictably.

3.3. Strange Attractors: Systems' disorder is channeled into patterns with common underlying themes. These "themes" attract or pull the normal energy flow; they are basins of attraction. Strange attractors enlarge the role of attractors; they excite and lure a system's vitality into non-projected patterns. Strange attractors create unpredictability and introduce data into the system. Strange attractors originate from varying initial conditions. Because of their own behavior they change "what we can know about them" (Hayles, 1990, p. 159). Examples would be the student question that throws the class discussion into a new direction, unseen but internal to the system.

Implications

Suggestions from chaos for social systems are beginning to emerge. Nevertheless, there are implications for the fields of communications and educational communications and technology. I will make the following observations:

Observation 1. In our daily activities, seemingly, humans and our organizations pass through processes similar to systems on their way to chaos. [A new idea, criticality (Bak and Chen, 1991; Peterson, 1989) may be important here. Criticality states that systems evolve to critical points, then often unknown and/or unplanned behaviors occur to set them in different directions.] Flach (1988), a psychiatrist, has a life cycle model that provides a parallel motif to chaos theory (Figure 1). He first notes that there are times in life when major shifts must occur. These shift points are bifurcations. Stress occurs to our resting system and we become severely destablized. Our own internal and or external structures may become chaotic. Such bifurcation points happen unexpectedly, as when very traumatic events take place, and expectedly, as when we pass through phases of our lives.

When our steady state systems are stressed, "disruption" takes place. Disruption will be filled with human and organizational turbulence. A chaotic state can follow wherein attractors may help "reintegration"--putting the disrupted parts together into a new homeostases. We self-organize into something unlike we were before. During the chaos, people and organizations cannot necessarily determine or choose future direction. Thus risk is present (Figure 2).

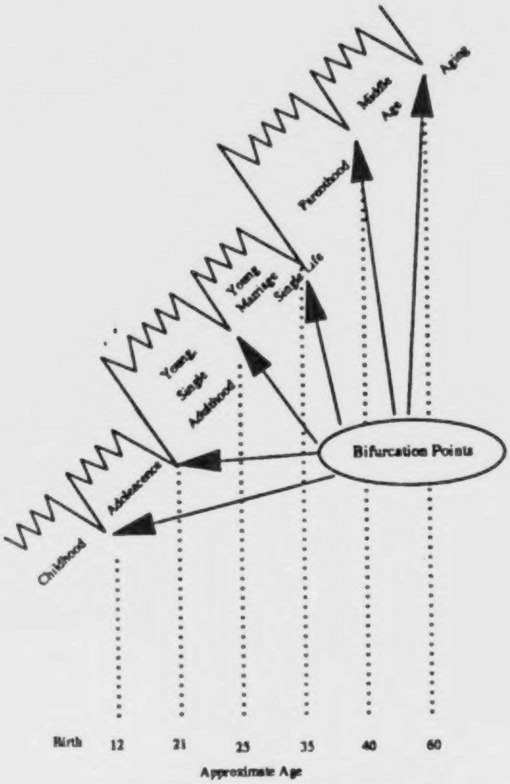


Figure 1: Life Cycle Stages (Flach, 1988)

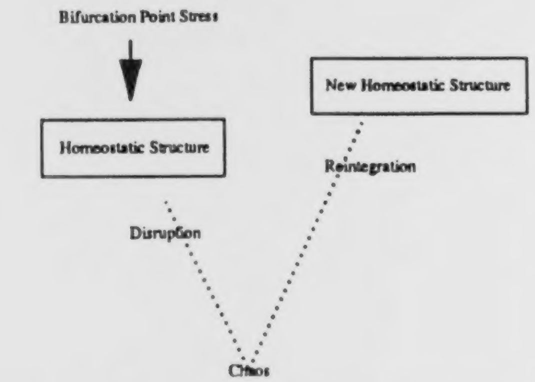

Bifurcation Point Stress
New Homeostatic Structure
Homeostatic Structure
Reintegration
Disruption
Chaos

Figure 2: The Normal Disruption-Reintegration Cycle (after Flach, 1988)

Flach (1988) also states that individuals may develop a negative homeostases as a future state. Figure 3 shows a failure to enter a disruptive state. Organizations, and I would add, ideas, can be seen to exhibit similar behavior under stress. A second type of negative outcome is shown below in Figure 4 (Flach, 1988). While disruptive behaviors occur, the system does enter an entropic phase where a dysfunctional outcome emerges.

Figure 3: The Failure to Disrupt (based on Flach, 1988)

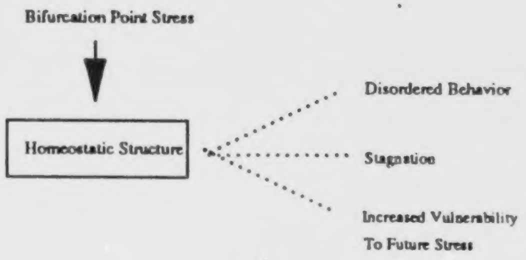
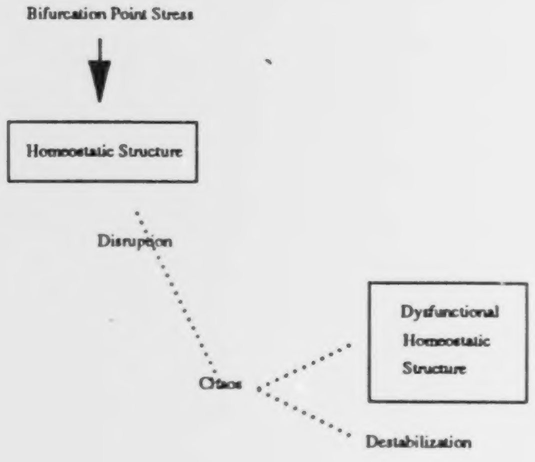

Bifurcation Point Stress
Homeostatic Structure
Disordered Behavior
Stagnation
Increased Vulnerability
To Future Stress

Figure 4: The Failure to Reintegrate Successfully (after Flach, 1988)


Bifurcation Point Stress
Homeostatic Structure
Disruption
Chaos
Dysfunctional
Homeostatic
Structure
Destabilization

An older model (Figure 5, Wheeler, et al. 1986) tried to visualize this idea. How each individual or organization responds to change depends on how change is perceived and also on the type of change occurring. Some changes are easy to accept and adapt to. Four types of change are:

a. Wanted/Expected: If we know something is going to occur and we want it, then the adjustment required is minimal. These types of changes include going to a meeting, the generation of a new idea, teaching, or buying new books.

b. Not wanted/Expected: If we know something is going to happen, but we do not want it to occur, the adjustment required is more taxing. An example of this might be our own physical changes as we age. We cannot avoid them and we usually do not want them to occur, but we cannot do anything about them. Other examples include paying Association dues, buying a full fare airplane ticket, or facing the death of a close teacher or mentor.

c. Wanted/Non Expected: Another type of change is one that we do not expect, but we are happy when it occurs. Such changes can still be stressful, but once the initial surprises and problems are resolved, adjusting to these changes is another pleasant assignment. Examples are unexpected visuals from colleagues, professors, or former students, receiving a promotion, or being invited to present at a large conference for a large fee.

d. Not Wanted/Not Expected: The last category of change is the most difficult to manage because we do not expect it and it is not desirable. Examples might include not getting tenure or promotion, total rejection of an article, or a break-in at the office.

	Wanted	Not Wanted
Expected	*Wanted* *Expected*	*Not Wanted* *Expected*
Not Expected	*Wanted* *Not Expected*	*Not Wanted* *Not Expected*

Figure 5: Four Types of Change

The implications here are: a) In people and human systems, bifurcation points can be both expected and unexpected, wanted and not wanted; b) The outcomes of these bifurcations are dependent on both initial conditions and attractors; c) Outcomes may be positive, when a new system emerges, or negative, where the current system falters, dysfunctions, regresses, or destablizes; d) The outcomes are relative. That is, depending on the position and values of the individual or system, outcomes may be seen to be positive or negative.

We should be aware of communication systems that are obstructionistic and debilitating (Figure 6: Flach, 1988). Such environments do not support positive systems and should be considered when implementing large scale instructional development projects. While not a series of guidelines, Figure 5 can serve as a checklist of design heuristics.

Figure 6. Debilitating Environmental Structures

Too Rigidly Organized	*Too Much Disarray*
Abusive or demanding	Indulgent
Excessive expectations	Murkey expectations
Stubborn resistance to change	Excessively fluid
Rejecting/ambivalent	Overly permissive
Discourages Communications	Indiscreet
Intolerant of conflict	Perpetuates turmoil
Punitive or vindictive	Excess forbearance
Thwarts hope	Pollyannaish
Rejects new ideas	Rejects tradition
Destructive values	Ill-defined values or none

Source: Flach, 1988, p. 216

In addition, Flach (1988) identified environmental elements that help develop new, positive systems. These are shown in Figure 7:

Figure 7: Elements for a positive environment

- coherent but flexible structures
- a human network
- respectfulness
- recognition
- assurance of privacy
- tolerance of change
- acceptance
- defined, realistic limits on behavior
- open communications
- responsiveness to new ideas
- tolerance of conflict
- promotion or reconciliation
- hopefulness
- a sense of community
- constructive human values
- empathy

(Flach, 1988, p. 213)

There is a simplicity and obviousness in Flach's work, which I believe assists communicators in examining and developing systems.

Observation 2. To examine certain complex ideas and systems, we may have to use and apply simpler systems (Pool, 1990b; Stewart, 1989). That means reviewing some of our more complicated and elaborate communication models and processes to identify their basic features. In these simpler systems, and in the simplicity of basic models, we may find answers to our communication and educational questions.

More complex things evolve from simpler things. But when complexity confuses issues and one needs a "whack on the side of the head," examining and studying simple systems may be helpful. Perceived complexity may mask obvious explanations. While not simplifying problems, simple models may "...yield extremely complex developments" (Baumol and Benhabib, 1989, p. 80). Stewart (1989) summarizes this point when he writes that "...apparently patternless behavior (sic) may become simple and comprehensible if you look at the right picture" (p. 47).

There is a major implication for our communication and educational technology curriculums (Francis and King, 1990). We need more process type learning activities. Formulas and facts, spoon fed, will not help anyone in the learning environments. Critical thinking, problem solving, creativity, and communication skills can, should, and must be part of each lesson and course. Books like *Drawing on the Right Side of the Brain* (Edwards, 1989, 2nd edition), *Exercises in Visual Thinking* (Wileman, 1980), *experiences in visual thinking* (sic) (McKim, 1972), *Use Both Sides of Your Brain* (Buzan, 1976), *The Art of Creative Thinking* (Nierenberg, 1982), *The All New Universal Travel* (Koberg and Bag, 1981, Rev. ed.), *Conceptual Blockbusting* (Adams, 1976), as well as others, must become part of the communication and educational technology curriculum.

Crutchfield, et al. (1986) write, "Even the process of intellectual progress relies on the injection of new ideas and on new ways to connecting old ideas. Innate creativity may have an underlying chaotic process that selectively amplifies small fluctuations and molds them into macroscopic coherent mental states that are experienced as thoughts. In some cases the thoughts may be decisions, or what are perceived to be the exercise of will. In this light, chaos provides a mechanism that allows for free will within in a world governed by deterministic laws" (p. 57).

Rapp (quoted in McAuliffe, 1990) discusses chaos activity as an asset in problem solving: "You want to be able to scan as wide a range of solutions as possible and avoid locking on to a suboptimal solution early on....One way to do that is to have a certain amount of disorderliness, or turbulence, in your search" (p. 48).

Observation 3. Sensitive dependence on initial conditions questions "whether an effect is proportional to a cause" (Hayles, 1990, p. 16). If initial conditions are known, it would seem that the educational systems evolution may be statistically predicted. Regardless, perfect knowledge of initial conditions is almost impossible. There will always be deviation. Because of the character of attractors, the status of the system at later times may be completely unlike the one planned. This has implications for both classroom and computer management and how teachers, designers, and administrators might approach such management. Dynamics of the instructional process and the interaction among the players are also affected. By more decidedly identifying initial conditions, chaos may be minimized (Curtis, 1990).

For example, chaos can characterize educational settings. Cziko (1989) notes that student achievement may be determined completely by the student's entry skills. The total achievement may be completely unpredictable. Cziko (1989) also points out that tiny knowledge differences between students at the start of instruction may lead to sizable and unpredictable differences between participants by the end of instruction. The lesson for researchers and practioners is that initial conditions have been, are, and will be critical in any educational setting.

Pool (1989a,b) summarizes that by saying that, because of chaos, researchers, at least, will now have an "appreciation" of how little complexity can produce complicated results.

While it is probably impossible to recognize and describe all initial conditions, there are many upfront tools to support communicators. For audience analysis, for example, the Myers-Briggs Temperament Inventory (Barrett, 1989; Barrett and Zamani, 1989; Keirsey and Bates, 1978; Kroeger and Thuesen, 1988; Lawrence, 1982, 2nd ed.) and the Learning Styles Inventory (Kolb, 1976, 1982) are useful instruments, and underexploited by educators. A future tool that may have a part to play in describing initial conditions is fuzzy set theory (Gui and Goering, 1990). This new area of mathematics deals with uncertainty due to lack of sharply defined boundaries. ("Big" is fuzzy; there are not distinctions or clear cut differentiations between "big" and "not big.") Through fuzzy set theory, ambiguous phenomena can be described. Other needed future tools also will include instruments to identify and describe communication sources more quickly, and mechanisms to provide a knowledge of the instructor and the instructor's teaching preference and style to both the instructor and the student.

Observation 4. Nature imposes restraints on planning. Chaos may be confined to areas or events that may be predictable, or at least anticipated in systems. Baumol and Benhabib (1989) note that "...where chaos occurs, extrapolation and estimation ...are questionable." This has implications for the flow of data in organizations. Another restraint is the attractor. What points exist, unplanned and unseen, to attract the system or its component elements?

We have to identify both initial and continuing patterns better. Use of communication case studies (King, 1984), in-class research, and qualitative descriptions (Jonassen, 1990) will have to be better incorporated into research paradigms, better supported by the research community, and better received by professional journals. Expanded use of Masters theses could be one way to carry out the research side.

Observation 5. We will have to recognize intrinsic, non-linear structures (Savit, 1990). However, with nonlinearity and chaos, there is one, degree of order, perhaps temporary, allowing predictability, anticipation, and management. Chaos says that "...apparently random behavior may not be random at all" (Baumol and Benhabib, 1989, p. 80). R. May (1989) wrote that "...we must expect to see chaos as often as we see cycles or steadiness...we would all be better off it more people realized that simple nonlinear systems do not necessarily possess simple dynamical properties" (p. 41).

The issue we face is to find the degree of order necessary to keep the ability, and the flexibility, to adapt, while using the variability inherent in a chaotic situation to advance, rather than continuing to use blinders or wringing our hands about the unfairness of it all (Francis, 1990). We will have to incorporate change in a "world dominated by interacting processes" (Francis, 1988). Systems analysis of communication systems will be rekindled (Saba, 1989).

Observation 6. We will have to look for scaling structures, examining, comparing, and relating big details to small details. The challenge here is how scale variance affects facts (Hayles, 1990, p. 16). Researchers will have to examine the variability of a system as well as its stable parts (Pool, 1989b,c,d). Individual situations or local systems that are perceived as unpredictable may have stable patterns when viewed from a suprasystems perspective (J. May, 1989). An in-class example might be cooperative learning, moving from the individual, to a group, to groups, to the class, to the entire school. Jonassen (1990) calls it "...seeking the correct scale for examining the learning process..." (p. 33). All organizations can be viewed through a scaling framework.

Observation 7. An issue researchers will have to face is figuring out what to measure. Sources of noise will have to be reviewed (Olsen and Schaffer, 1990). When "noise" is found in a system, and when random behaviors occur, researchers will have to deduce the "attractors" that cause such activity. Distinguishing between chaos and noise will require technical calculations of the data (Pool, 1989a,c,d,e,f) using high level mathematics, powerful computers, and programming, and is both difficult (Meyer, 1991) and "...in its infancy..." (R. May, 1989).

Combining observations 6 and 7, researchers also will need to examine the scale from which to examine. Looking at systems, Paulre (1981) first identifies what to study -- three categories of examination:

a. explanation of an observed changing system,
b. analysis of conditions and effects of change, and
c. forecasting or conjecture of change.

He then discusses scaling, from local, to global, to normal organizational analysis. Paulre developed this mix of scale and components to measure into a 3x3, definition and classification matrix (Figure 8). Such a matrix may offer some researchers a framework for communication analysis, combining scale and content.

	Explanation of an observed changing system	Analysis of conditions and effects of change	Forecasting or Conjecture of change
Local Analysis			
Global Analysis; Phasing & cycles			
Organizational System Structure			

Figure 8: A Framework to define and classify systems (Paulre, 1981)

In examining what to measure, questions are raised about research problems in general. Fuchs and Fuchs (1990) discuss problems with educational research that are reflected in the communication research. These problems include, first, an emphasis on research that is theory driven and lab based rather than program building research, and second, research that is viewed as unimportant, irrelevant, too fractionalized, directionless, and finally, uninterpretable to many. Solutions include in-class communication research as well as participatory research with the teachers and local communication people as team members.

Observation 8. Fractals have an interesting potential for graphic designers (Barnsley, 1988;Barnsley, et al., 1988; Pool, 1989c; Pool, 1990a). The beauty of the fractal patterns seems to have visual power. These images capture attention and command respect. Mandelbrot (1989, 1990) feels that the visual nature of fractals has made them an area of interest. Apparently the mathematics of fractals will soon be used to compress and reconstruct images for digitized visuals.

Conclusions

Chaos offers educational communications and technology new systems' tools and ideas. Chaos provides "the system with a deterministic "I don't know state" within which new activity patterns can be generated.

Previous characterizations and paradigms of systems stated:

1. simple systems behaved in simple ways,
2. complex behaviors implied complex causes, and
3. different systems behaved differently. (Gleick, 1987)

These ideas are now undergoing serious and exciting evaluation and re-examination because of chaos. This shift is producing new ways of studying problems:

1. simple systems give rise to complex behavior and involved descriptions,
2. complex systems may be studied by looking at simple behaviors and simpler systems. (Baumol and Benhabib, 1989; Gleick, 1987), and
3. much complexity holds universally, unrelated to the type of system, or its components.

Summarizing the exploding transformations based on systems and chaos, Linstone (1989) writes that now we have:

1. new ways to study complex systems,
2. techniques to study problems using multiple perspectives, and
3. approaches to employ "...'what if' rather than prediction as a basis for planning" (p. 12).

Prigogine and Stengers (1984) assert that systems are "undergoing a radical change toward the multiple, the temporal, and the complex" (p. xxvii). Chaos is a toolbox of ideas (and statistical methods) that can aid us by providing communicators with new insights, multiple perspectives, and "what if...?" attitudes.

REFERENCES

Ackoff, R. L. 1971. Towards a system of systems concepts. Management Science. 17(11): 661-671.

Adams, J. L. 1978. Conceptual Blockbusting. New York, NY: W. W. Norton.

Ashmos, D. P. and G. P. Huber. 1987. The systems paradigm in organization theory. Academny of Management Review. 12(4): 607-621.

Bak, P. and K. Chen. 1991. Self-organized criticality. Scientific American. 264(1): 46-53.

Barnsley, M. 1988. Fractals Everywhere. San Diego, CA: Academic Press, Inc.

Barnsley, M. F., A. Jacquin, F. Malassenet, L. Reuter, and A. D. Sloan. 1988. Harnessing chaos for image synthesis. ACM Computer Graphics. 22(4): 131-138.

Barrett, L. A. 1989. Impact of teacher personality on classroom environment. J. Psychological Type. 18(Winter): 50-56.

Barrett, L. A. and H. Zamani. 1989. Learning strategies that agricultural students use successfully and unsuccessfully. Paper presented at the National Agricultural Research Meeting. Orlando, FL. December.

Baumol, W. J. and J. Benhabib. 1989. Chaos: significance, mechanism, and economic applications. J. Economic Perspectives. 3(1): 77-105.

Boulding, K. E. 1956. General systems theory--The skeleton of science. Management Science. 2(3): 197-208.

Buzan, T. 1976. Use Both Sides of Your Brain. New York, NY: E. P. Dutton.

Chaitin, G. 1990. A random walk in arithmetic. New Scientist. 25(1709): 44-46.

Crutchfield, J. P., J. D. Farmer, N. H. Packard, and R. S. Shaw. 1986. Chaos. Scientific American. 255(6): 46-57.

Cziko, G. A. 1989. Unpredictability and indeterminism in human behavior: arguments and implications for educational research. Educational Researcher. 18(3): 17-25.

Curtis, R. K. 1990. Complexity & predictability: The application of chaos theory to economic forecasting. Future Research Q. 6(4): 57-70.

Dervin, B. 1989. Users as research inventions. J. Communication. 39(3): 216-232.

Dunbar, S. R. 1990. A tourist guide to chaos and fractals. Presented at the Symposium on Chaos in Biological and Agricultural Systems: Statistical Issues. Lincoln, NE, June 3-5.

Edwards, B. 1989. Drawing on the Right Side of the Brain (2nd ed.). Los Angeles, CA: Tarcher.

Flach, F., M.D. 1988. Resilience: Discovering a New Strength at Times of Stress. New York, NY: Fawcett Columbine.

Francis, C. A. 1988. Unique importance of information in future agricultural production systems. Personal Communication.

Francis, C. and J. King. 1990. Preparing Nebraska students for lifelong learning in agriculture. Crop Production News. 10(25). 2pgs.

Fuchs, D., and L. S. Fuchs. 1990. Making educational research more important. Exceptional Children. 57(2): 102-107.

Gharajedaghi, J., and R. Ackoff. 1984. Mechanisms, organisms and social systems. Strategic Management J. 5: 289-300.

Glass, L., and M. C. Mackey. 1988. From Clocks to Chaos: The Rhythms of Life. Princeton, NJ: Princeton University Press.

Gleick, J. 1987. Chaos: Making a New Science. New York, NY: Viking Penguin, Inc.

Gordon, T. J., and D. Greenspan. 1988. Chaos and fractals: new tools for technological and social forecasting. Technological Forecasting and Social Change. 34: 1-25.

Gui, X. Q., and C. E. Goering. 1990. Introduction to fuzzy set theory and applications. Transactions of the ASAE. 33(1): 306-313.

Hayles, N. K. 1990. Chaos Bound: Orderly Disorder in Contemporary Literature and Science. Ithaca: Cornell University Press.

Heppenheimer, T. A. 1989. Classical mechanics, quantum mechanics, and the arrow of time. MOSAIC. 20(2): 2-11.

Huberman, B. A. 1989. An ecology of machines: how chaos arises in computer networks. The Sciences. 29(4): 38-44.

Jonassen, D. H. 1990. Thinking technology: chaos in instructional design. Educational Technology. 30(2):32-34.

Keirsey D. and M. Bates. 1978. Please Understand Me. Del Mar, CA: Prometheus Nemesis.

King, J. W. 1984. A framework for using case study methodology to describe development communication programs. DevCom Q. 2(3): 1-11.

Koberg, D. and J. Bagnall. 1981. The All New Universal Traveler (rev. ed.). Los Altos, CA: William Kaufmann.

Kolb, D. A. 1976. Learning Style Inventory: Technical Manual. Boston, MA: McBer.

Kolb, D. A. 1982. Personal Learning Guide: Leader Notes. Boston, MA: McBer.

Kroeger, O. and J. M. Thuesen. 1988. Type Talk. New York, NY: Tilden Press.

Lawrence, G. 1986. People Types and Tiger Stripes (2nd ed.). Gainesville, FL: Center for Applications of Psychological Type.

Lehrer, R., R. C. Serlin, and R. Amundson. 1990. Knowledge or certainty? A reply to Cziko. Educational Researcher. 19(6): 16-19.

Mandelbrot, B. 1990. Fractals--a geometry of nature. New Scientist. 127(1734): 38-43.

Mandelbrot, B. B. 1989. Fractals and an art for the sake of science. LEONARDO. Computer Art in Context Supplemental Issue. 21-24.

May, J. H. 1989. Chaos and the marketing of computing services on campus. CAUSE/EFFECT. 12(1):4-6.

May, R. M. 1989. The chaotic rhythms of life. New Scientist. 124(1691): 37-41.

May, R. M. 1976. Simple mathematical models with very complicated dynamics. Nature. 261: 459-467.

McAuliffe, K. 1990. Get smart: controlling chaos. OMNI. 12(5): 42-48, 86-92.

McKim, R. H. 1972. Experiences in Visual Thinking. Montery, CA: Brooks/Cole.

McRobie A., and M. Thompson. 1990. Chaos, catastrophes and engineering. New Scientist. 126(1720): 41-46.

Meyers, G. 1991. Personal Communication.

Nierenberg, G. I. 1982. The Art of Creative Thinking. New York, NY: Cornerstone.

Olsen, L. F., and W. M. Schaffer. 1990. Chaos versus noisy periodicity. Science. 249: 499-504.

Paulre, E. 1981. System Dynamics and Analysis of Change. New York: North-Holland Publishing Company.

Peterson, I. 1989. Digging into sand. Science News. 136(3): 4042.

Pool, R. 1989a. Chaos theory: how big an advance? Science. 245:26-28.

Pool, R. 1989b. Ecologists flirt with chaos. Science. 243:310-313.

Pool, R. 1989c. Is it chaos, or is it just noise? Science. 243:25-28.

Pool, R. 1989d. Is it healthy to be chaotic? Science. 243:604-607.

Pool, R. 1989e. Is something strange about the weather? 243:1290-1293.

Pool, R. 1989f. Quantum chaos: enigma wrapped in a mystery. 243:893-895.

Pool, R. 1990a. Fractal fracas. Science. 249:363-364.

Pool, R. 1990b. Putting chaos to work. Science. 250:626-627.

Prigogine, I. and I. Stengers. 1984. Order Out of Chaos. New York, NY: Bantam.

Puccia, C. J., and R. Levins. 1985. Qualitative Modeling of Complex Systems: An Introduction to Loop Analysis and Time Averaging. Ch. 1. Qualitative Models. p. 1-11. Cambridge, Massachusetts: Harvard University Press.

Saba, F. 1989. Personal Communication.

Salisbury, D. F. 1990. General systems Theory and instructional systems design. Performance and instruction. 29(2): 1-11.

Saperstein, A. M. 1984. Chaos--a model for the outbreak of war. Nature. 309: 303-304.

Savit, R. 1990. Chaos on the trading floor. New Scientist. 127(1729):48-51.

Shannon, C. E., and W. Weaver. 1949. The Mathematical Theory of Communication. Urbana: University of Illinois Press.

Shaw, R. 1984. The Dripping Faucet as a Model Chaotic System. Santa Cruz, CA: Aerial Press, Inc.

Shaw, R. 1981. Strange attractors, chaotic behavior, and information flow. Z. Naturforsch. 36a: 80-112.

Stewart, I. 1989. Portraits of chaos. New Scientist. 124(1689): 42-47.

Toffler, A. 1984. Science and change. In I. Prigogine and A. Stengers. 1984. p. xi-xxvi.

Tsonis, A. A., and J. B. Elsner. 1989. Chaos, strange attractors, and weather. Bulletin American Meteorological Society. 70(1): 14-23.

von Oech, R. 1988. Creative Whack Pack. Stamford, CT: U.S. Games Systems.

Wheeler, D. W., J. W. King, H. G. Lingren, D. W. Swoboda, and J. Randall. 1986. Helping Rural People Under Stress. Lincoln, NE: Nebraska Cooperative Extension Service.

Wileman, R. E. 1980. Exercises in Visual Thinking. New York, NY: Hastings House.

Winn, W. 1990. Some implications of cognitive theory for instructional design. Instructional Science. 19(1): 53-69.

Title:

A Fine-tuned Look at White Space Variation in Desktop Publishing

Authors:

Nancy Nelson Knupfer
Marina Stock McIsaac

A Fine-Tuned Look at White Space Variation in Desktop Publishing

By
Nancy Nelson Knupfer, Ph.D. and Marina Stock McIsaac, Ph.D.
Arizona State University

Desktop publishing has revolutionized the publication industry. Layout programs permit designers to choose various layout styles; place text and graphics within pages; decide on fonts, hyphenation, justification; and select color separations. Three of the most popular programs, QuarkXPress, Pagemaker , and Ready, Set Go! accept styling information from imported manuscripts and allow an author to write using character formatting, kerning, bold-face, italic, fonts, and line-break information. Clip art, photographs, technical illustrations, and so forth can be inserted into a variety of text formats. Three sophisticated computerized text formats which have been made available recently to consumers, and are used to market the desktop publishing software, include run-around text, wrap-around text, and transparent text.

Run-around text is boxed around the graphic and has straight borders (see Figure 1). Wrap-around text conforms to the shape of the graphic and has irregular borders (see Figure 2.) And, transparent text is that which has a graphic placed directly over the text (see Figure 3). In all three of the top programs, the software allows the user to place text over a graphic and vice versa. It also allows the user to run text around a rectangular graphic box, run text around the outline of a graphic, or run text behind a graphic box. Examples of all three formats can be found in consumer magazines, newspapers, and instructional materials.

[Insert Figures 1, 2, and 3 about here.]

As a result of new publishing enhancements, people with little publishing experience are now making decisions involving the flow of text on pages, placement of graphics within text, and text wrap-around or run-around features. Such decisions, often made with no guiding principles in mind, affect both reading speed and comprehension (Knupfer & McIsaac, 1989a). In many cases, inappropriate selections are being made which adversely affect readability. If we can examine the effects of many of these page-layout enhancements and develop principles for their effective use, perhaps these design choices can become more educationally sound. The amount of white space surrounding a graphic (Knupfer & McIsaac, 1990a) may well be one of those variables which can be manipulated to enhance understanding of the text .

BACKGROUND

Text variables and graphic design have received attention in four basic areas of research: graphic design, instructional text design, instructional graphics, and, in recent years, computer design of materials for on-screen or printed material.

Research on graphic design emphasizes the importance of balancing text with white space, improving the aesthetics of the page, and positioning graphics as the dominant visual element (Parker, 1987). Such research does not usually question the content of the graphic. And, research in the area of reading indicates that the type of graphic is an important variable in reading comprehension, but that most students don't attend to the graphics (Reinking, 1986). Simple graphics at the pictorial level might not make a difference in reading speed or comprehension, but graphics which are designed at the inferential, generalized, and evaluative levels of understanding can help students grasp the meaning of the text (Reinking, 1986; Singer & Donlan, 1980). Further, research on learning suggests that designers should address cognitive processes by developing materials for graphic thinkers, not just graphic readers (Boyle, 1986).

Recently, principles for text design have been developed for use by instructional designers (Hartley, 1985; Heines, 1984; Jonassen, 1982). Layout is one variable which significantly affects readers' speed and comprehension. But the layout variables in text research (type style, upper-lower case, character attributes, line length, and justification) do not address the question of graphic content or placement. There is a dearth of reported research on graphic placement in instructional text which can influence guidelines for the emerging capabilities of new computer software programs (Isaacs, 1987; Knupfer & McIsaac, 1990a).

Instructional graphics research reveals no empirical evidence that the use of graphics in instruction improves learning, but the graphics used in that research might have represented concepts already familiar to the readers (Merrill & Bunderson, 1981). More sophisticated instructional graphics, such as pictorial, symbolic, schematic, or figurative displays, can be effective in teaching, and the visualization of abstract ideas through figurative displays may very well enhance learning (Nygard & Ranganathan, 1983).

Research that compared reading comprehension using computer screen versus printed text revealed nonsignificant differences (Gambrell, Bradley, & McLaughlin, 1987). Despite the similar result concerning comprehension, students reported more difficulty reading from the computer screen than from printed text; one variable which caused this difficulty was fatigue (Hathaway, 1984; Mourant, Lakshmanan, & Chantadisai, 1981). Even though research on electronic and printed-based materials have their own specific variables, certain features apply to both, including text density, upper- and lower-case, and typographic cueing (Hartley, 1987; Hathaway, 1984; Morrison, Ross, & O'Dell, 1988; Ross, Morrison, & O'Dell, 1988).

Several researchers have investigated layout, typographic cueing, and representation of graphic material in print-based text. Hartley (1985) determined that page size affects other layout decisions. Some studies that have compared the effects of justified and unjustified text on reading speed and comprehension found no significant differences (Campbell, Marchetti, & Mewhort, 1981; Muncer, Gorman, Gorman, & Bibel, 1986), yet the issue remains unresolved (Jonassen, 1982). Jandreau, Muncer, and Bever (1986) have investigated the effects of varying spaces between phrases in sentences. Still others have examined certain design features common to printed and electronic text in the presentation of graphic materials (Alesandrini, 1987; Poggenpohl,1985; Winn,1987; Winn & Holliday, 1982). Hartley (1987) and others have requested that such research be done.

Knupfer and McIsaac (1990a, 1990b, 1989a, 1989b) investigated the effects of various wrap-around enhancements which were produced electronically for print-based materials. They found that the wrap-around text placement produces faster reading time and better comprehension than the run-around text placement (1989a). Further, when comparing the use of the wrap-around text format, with variations in the amount of white space between the graphic and the text, they found a trend but no significant differences in reading speed or comprehension between text that was placed very close to the graphic and text that was placed further away from the graphic, showing 1/4 inch of white space (1990a).

Although certain features of electronic text can be transferred to printed text, there are many differences between the two which suggest that separate studies be done. The size of the screen or page, the type size and the ability to scroll are but a few of the differences which affect print and electronic text comparison studies. Because of these differences, a programmatic series of investigations is needed to examine the effects of these recent technological enhancements in each of the two media formats. Recent work by Morrison, Ross, and O'Dell (1988) examines attributes of computer based instruction, and attempts to affect these attributes by regulating text density for print and screen display lessons. Results suggested that print had more advantages than the computer screen in achievement, completion time, and learning efficiency for a statistics lesson. Therefore, the present study is limited to the examination of print-based, computer-generated text.

PURPOSE

Graphic designers have long emphasized the importance of using white space for balancing text, positioning graphics, and improving the aesthetics of the page layout (Parker, 1987). However, few attempts have been made to examine how the shape and amount of white space between the graphic and text enhances the visual without adversely affecting comprehension of the reading material.

The purpose of this study was to investigate previously noted trends (Knupfer & McIsaac, 1990a, 1990b, 1989a, 1989b) in order to discover the point at which the white space interferes with reading speed and comprehension. This study utilized only the wrap-around text format and compared the effects of variations in white space surrounding the graphic. Wrap-around text with no white space around the graphic (see Figure 2) was compared to wrap-around text with one-half inch of white space between the graphic and text (see Figure 4). A third group which contained no graphic at all was used as a control (see Figure 5).

[Insert Figures 4 and 5 about here.]

Based on the principles of perceptual organization, which include similarity, proximity, continuity, and closure (Bloomer, 1976) we speculated that the amount of white space surrounding the graphic would affect reading reading speed and comprehension. These four processes, by which the mind organizes meaning, depend on how physically close the objects are, how similar they are, whether there is a continuous line to guide the eye, and whether the minimal amount of information is present that is necessary to obtain meaning or closure. Reading speed and comprehension are directly affected by the way the mind organizes meaning from the placement of graphics and text. In particular, the principle of proximity states that the closer two or more visual elements are, the greater is the probability that they will be seen as a group or a pattern. Further, the critical nature of horizontal space between visual elements influences the interpretation of the meaning of those elements (Zakia, 1975). Based upon this information, this study tested the hypothesis that reading speed and comprehension would be significantly greater when text was wrapped tightly around the graphic than when it had one-half inch of white space surrounding it.

METHOD

Subjects

One hundred, twenty-eight students from an introductory computer literacy course at a major university in the Southwest participated in this study. The elective course was an upper division course which was offered to meet a General Studies numeracy requirement. Students were sophomores, juniors, and seniors who had very little or no prior computer experience. This study was conducted at the beginning of the semester before the students encountered the information presented in the textual material. All students in the course participated in the study.

Materials

Readings. Readings included approximately six pages of informative text about Hypercard and desktop publishing. These materials were chosen because the topic was relevant to the subjects and because the reading level was geared for the general public as opposed to technical readers. The topic was suited to a novice audience with no prior knowledge about either Hypercard or desktop publishing. The text was analyzed using the Flesch-Kincaid Readability formula and was found suitable for college students between the 11th and 13th grade levels.

Three sets of reading materials were used for this study. All included the same text arranged in single column format. Set one did not contain any graphics at all. Sets two and three each contained 12 identical graphics which were placed in the text using the wrap-around technique. The only difference between sets two and three was the amount of white space surrounding the graphics. Set two had no white space between the graphics and text, while set three had one-half inch of white space between the graphics and text. The simple, pictorial graphics were meant to be motivational and were linked to the text by interest. Even though they were not critical for the understanding of the textual material, the graphics very well could have promoted understanding through illustrative associations.

In the hypercard article, the graphics included a Macintosh computer with a set of linked cards on the screen corresponding to an introductory paragraph describing hypercard and the machinery it runs on, a stack of index cards representing Hypercard's index card analogy, a set of books representing computer manuals, an organizational chart representing the five levels of hypercard application, an educator sitting at a computer representing Hypercard's impact on education, a teacher and an administrator discussing the potential of Hypercard in schools, a set of computer disks representing Hypercard's impact on new software development, and finally, a newsletter and newspaper set representing the important evolutionary step in computing attributed to Hypercard. The desktop publishing article displayed a desk to represent the desktop publishing concept, a hand on a mouse where the text explains that various desktop publishing programs support mouse usage, and a set of black and white tack pins representing the concept of the same yet different, to coincide with the text's explanation concerning variations in the same desktop publishing software to accommodate the MS-DOS and Macintosh environments.

Although some professional layout designers might question the graphics placement within these reading materials in terms of the elements of good design, it is important to know that the way in which the materials were developed is indicative of common practice among lay people. Therefore, any flaws in the materials design are potential problems that other teachers and trainers might exhibit as well.

Pretest and Posttest. To measure the students' comprehension of the reading material, a pretest and a posttest were developed from key concepts which the authors extracted from the readings. No test items were dependent upon the information contained in the graphics. The tests were developed and validated during a previous phase of the project, and were found to be reliable at the .76 level using the KR20 formula. The pretest and posttest were nearly identical; each contained the same 34 multiple choice and 11 completion questions utilizing the fill-in-the-blank format. The two types of questions measured both recognition and recall comprehension skills. In addition to those questions, the two groups with graphics were asked to recall as many graphics as possible on the posttest. They were told that there were 12 graphics to recall.

Procedure

All subjects were given a pretest to determine their knowledge of the subject prior to reading. There was no time limit and students were free to take the time they needed to complete the test. Next, the three sets of reading materials were randomly distributed to the students so that each student received one set of articles. Students were instructed to read through the text one time without hurrying. When subjects were finished reading, each of them recorded their time from a master timekeeper which posted times to the half second. Proctors observed the students to make sure that their times were recorded correctly. Immediately following the reading, the subjects were given the posttest. They were allowed as much time as they needed to complete the test.

Analysis

Even though the treatment groups were randomly chosen, there were slight variations in the mean pretest scores which had potential influence on the remaining statistical analysis. To determine appropriate statistical tests, analysis of variance (ANOVA) was conducted on the pretest scores for the three groups to determine if there were any significant differences in the groups prior to treatment. Independent variables were the text-graphic format (no graphics, wrap-around text with no white space, and wrap-around text with one-half inch of white space). The dependent variables included reading speed and comprehension. Since there were no significant differences in mean pretest scores for the three treatment groups (see Table 1), ANOVA was also conducted on the posttest scores to measure differences in comprehension. ANOVA was also conducted on the reading speed for the three groups. Additionally, a t-test was conducted on the graphic recall question for groups two and three, the two groups which had graphics. For all statistical testing, the alpha level for significance was set at .05.

TABLE 1
Summary of ANOVA
Pretest Comprehension by Format

Source	D.F.	Sum of Sq.	Mean Sq.	F Ratio	F Prob.
Between Groups	2	69.23	34.61	.84	.43*
Within Groups	125	5162.65	41.30		
Total	127	5231.88			

Not significant at the .05 level

RESULTS AND CONCLUSIONS

Significant differences in comprehension were found between the groups. Subjects in group three, with one-half inch of white space between the graphic and text, comprehended significantly less than subjects in group one and group two; groups one had no graphics and group two had no white space between the graphic and text. There were no significant differences in reading speed between the three groups. Mean scores and standard deviations for the pretest, posttest, and reading speed for each of the treatment groups are shown in Table 2. Table 3 shows results of the ANOVA conducted on reading speed and Table 4 reports the ANOVA for posttest comprehension between treatment groups.

TABLE 2
Reading Pretest, Posttest, and Reading Speed in Minutes
Means and Standard Deviations

		Pretest		Posttest		Minutes	
Group	n	M	SD	M	SD	M	SD
No Graphic	44	7.09	6.15	40.14	12.08	9.86	1.98
Wrap Tight	44	6.04	5.79	38.23	15.17	10.54	1.98
Wrap 1/2" White	40	7.85	7.32	32.25	12.32	10.27	2.24

TABLE 3
Summary of ANOVA
Reading Speed by Format

Source	D.F.	Sum of Sq.	Mean Sq.	F Ratio	F Prob.
Between Groups	2	10.36	5.18	1.21	.30*
Within Groups	125	534.57	4.28		
Total	127	544.93			

Not significant at the .05 level

TABLE 4
Summary of ANOVA
Posttest Comprehension by Format

Source	D.F.	Sum of Sq.	Mean Sq.	F Ratio	F Prob.
Between Groups	2	1401.56	700.78	3.96	.02*
Within Groups	125	22094.41	176.76		
Total	127	23495.97			

Significant at the .05 level

We venture to guess that the lower mean comprehension score on for group three reflects the principal of visual perception called proximity; it is more difficult for the eye to follow the text and perceive meaning due the larger gaps that the eye must jump across to continue reading. Further, it agrees with Zakia's (1975) claim that the horizontal spacing is critical to accurate perception.

Table 5 shows results of the t-test between groups two and three concerning recall of the 12 graphics. There were no significant differences between the two groups based upon the amount of white space surrounding the graphic. The low mean scores indicate that neither group remembered very many of the graphics. This, along with similar reading speed between the groups, supports Reinking's claim that readers do not really pay attention to pictorial graphics (1986).

TABLE 5
T-Test
Graphics Recall by Format

Group	n	Mean	SD	SE	F Value	2-Tail Prob.
Wrap Tight	44	1.27	1.66	0.25	1.17	0.61*
Wrap 1/2" White	40	1.57	1.53	0.24		

Not significant at the .05 level

Based upon the findings of this research and our past research on this topic (Knupfer & McIsaac, 1990a, 1990b, 1989a, 1989b), we recommend that people who place graphics within text using desktop publishing software either use no graphics at all or use graphics which use the wrap-around text as opposed to the run-around, boxed style of text wrap. Further, when using the wrap-around text format, limit the white space between the graphic and text to less than one-half inch.

It is possible that reading speed and comprehension using these text-wrap configurations could vary dependent upon the type of graphic used. Therefore, we suggest that future research be conducted using a variety of graphics, including those which are more critical to the meaning of the text.

REFERENCES

Alesandrini, K.L. (1987). Computer graphics in learning and instruction. In H.A. Houghton and D.M. Willows (Eds.), *The psychology of illustration: Vol 2: Instructional issues*, New York: Springer Verlag.

Bloomer, C.M. (1976). *Principles of Visual Perception*. New York: Reinhold.

Boyle, M.W. (1986). Hemispheric Laterality as a Basis of Learning: What We Know and Don't Know. In G.D. Phye and T. Andre (Eds.), *Cognitive Classroom Learning: Understanding , Thinking, and Problem Solving*. San Diego, CA: Academic Press, Inc.

Campbell, A.J., Marchetti, F.M., & Mewhort, D.J.K. (1981). Reading speed and text production: A note on right-justification techniques. *Ergonomics*, 24, 611-622.

Gambrell, L.B., Bradley, V.N., & McLaughlin, E.M. (1987). Young Children's Comprehension and Recall of Computer Screen Displayed Text. Journal of Research in Reading, 10(2), 156-163.

Hathaway, M.D. (1984, January). Variables of Computer Screen Display and How they Affect Learning. *Educational Technology*, 7-11.

Hartley, J. (1987). Designing electronic text: The role of print-based research. *Educational Communication and Technology*, 35(1), 3-17.

Hartley, J. (1985). *Designing instructional text* (2nd ed.). New York: Nichols.

Heines, J. (1984). *Screen Design Strategies for Computer Assisted Instruction*. Bedford, MA: Digital Press.

Jandreau, S.M., Muncer, S.J., & Bever, T.G. (1986). Improving the readability of text with automatic phrase-sensitive formatting. *British Journal of Educational Technology*, 17(2), 128-133.

Jonassen, D.H. (1982). *The Technology of Text*. Englewood Cliffs, NJ: Educational Technology Publications.

Knupfer, N.N., & McIsaac, M.S. (1990). The Effects of Varying White Space Around Graphics Within Text. *Journal of Visual Literacy*. 10(1), 63-74.

Knupfer, N.N., & McIsaac, M.S. (1990). Visual Discrimination, Learning, and Page Layout: Wrap-Around, Run-Around, and Transparent Text. *Proceedings of the annual conference of the Association of Educational Communications and Technology* (AECT), Anaheim, CA. Jan. 30 - Feb. 4.

Knupfer, N.N., & McIsaac, M.S. (1989). Desktop Publishing: The Effects of Computerized Formats on Reading Speed and Comprehension. *Journal of Research on Computing in Education*. 22(2), 127-136).

Knupfer, N.N., & McIsaac, M.S. (1989). White Space, Graphics, and Text.. *Proceedings of the 21st Annual Conference of the International Visual Literacy Association* (IVLA). Scottsdale, AZ. Oct. 1989.

Merrill, P.F., & Bunderson, C.V. (1981). Preliminary Guidelines for Employing Graphics in Instruction. *Journal of Instructional Development*, 4(4), 1-9.

Morrison, G.R., Ross, S.M., & O'Dell, J.K. (1988). Text density Level as a Design Variable in Instructional Displays. *Educational Communication and Technology*, 36(1), 103-115.

Mourant, S.J., Lakshmanan, R., & Chantadisai, R. (1981). Visual Fatigue and Cathode Ray Tube Display Terminals. *Human Factors*, 23, 529-540.

Muncer, S.J., Gorman, B.S., Gorman, S., & Bibel, D. (1986). Right is Wrong: An Examination of the Effect of Right Justification on Reading. *British Journal of Educational Technology*, 17(1), 5-10.

Nygard, K.E., & Ranganathan, B. (1983, Spring). A System for Generating Instructional Computer Graphics. *AEDS Journal*, 16(3), 177-187.

Parker, R.C. (1987). *The Aldus guide to basic design.* Seattle: Aldus Corporation.

Poggenpohl, S. (Ed.) (1985). Graphic design: Computer graphics [Special issue]. *Visible Language*, XIX(2), 173-306.

Reinking, D. (1986). Integrating Graphic Aids into Content Area Instruction: The Graphic Information Lesson. *Journal of Reading*, 30(2), 146-151.

Ross, S.M., Morrison, G.R., & O'Dell, J.K. (1988). Obtaining more out of less text in CBI: Effects of varied text density levels as a function of learner characteristics and control strategy. *Educational Communications and Technology Journal*, 36(3), 131-142.

Singer, H., & Donlan, D. (1980). *Reading and Learning from Text.* Boston: Little, Brown.

Winn, W. (1987). Using charts, graphs and diagrams in educational materials. In D. M. Willows and H.A. Houghton (Eds.) *The psychology of illustration: Vol.1. Basic Research.* New York: Springer Verlag.

Winn, W.D., & Holliday, W.G. (1982). Design principles for diagrams and charts,In D.H. Jonassen (Ed.), *The technology of text.* Englewood Cliffs, NJ: Educational Technology Publications.

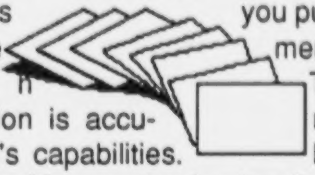
Zakia, R.D. (1975). *Perception and Photography.* Englewood Cliffs, NJ: Prentice Hall.

FIGURE 1

Run-around Text

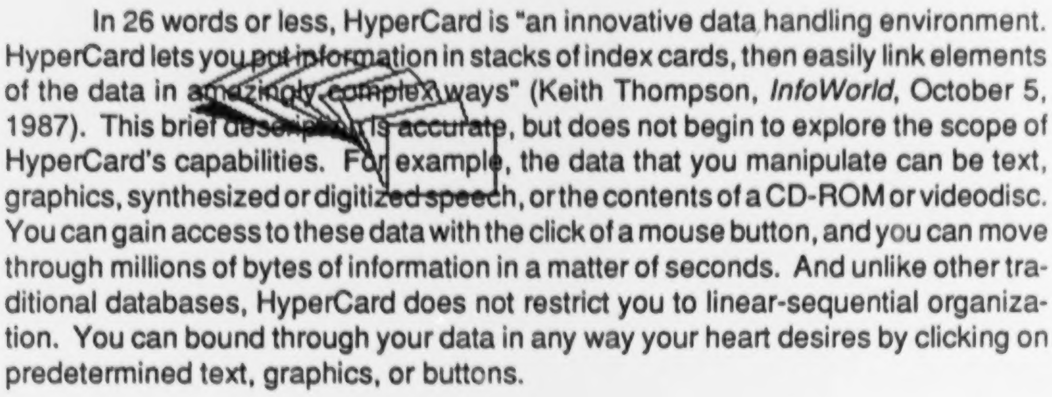
In 26 words or less, HyperCard is "an innovative data handling environment. HyperCard lets you put information in stacks of index cards, then easily link elements of the data in amazingly complex ways" (Keith Thompson, *InfoWorld*, October 5, 1987). This brief description is accurate, but does not begin to explore the scope of the HyperCard's capabilities. For example, the data that you manipulate can be text, graphics, synthesized or digitized speech, or the contents of a CD-ROM or videodisc. You can gain access to these data with the click of a mouse button, and you can move through millions of bytes of information in a matter of seconds. And unlike other traditional databases, HyperCard does not restrict you to linear-sequential organization. You can bound through your data in any way your heart desires by clicking on predetermined text, graphics, or buttons.

FIGURE 2
Wrap-around Text



In 26 words or less, HyperCard is "an innovative data handling environment. HyperCard lets you put information in stacks of index cards, then easily link elements of the data in amazingly complex ways" (Keith Thompson, *InfoWorld*, October 5, 1987). This brief description is accurate, but does not begin to explore the scope of HyperCard's capabilities. For example, the data that you manipulate can be text, graphics, synthesized or digitized speech, or the contents of a CD-ROM or videodisc. You can gain access to these data with the click of a mouse button, and you can move through millions of bytes of information in a matter of seconds. And unlike other traditional databases, HyperCard does not restrict you to linear-sequential organization. You can bound through your data in any way your heart desires by clicking on predetermined text, graphics, or buttons.

FIGURE 3
Transparent Text



In 26 words or less, HyperCard is "an innovative data handling environment. HyperCard lets you put information in stacks of index cards, then easily link elements of the data in amazingly complex ways" (Keith Thompson, *InfoWorld*, October 5, 1987). This brief description is accurate, but does not begin to explore the scope of HyperCard's capabilities. For example, the data that you manipulate can be text, graphics, synthesized or digitized speech, or the contents of a CD-ROM or videodisc. You can gain access to these data with the click of a mouse button, and you can move through millions of bytes of information in a matter of seconds. And unlike other traditional databases, HyperCard does not restrict you to linear-sequential organization. You can bound through your data in any way your heart desires by clicking on predetermined text, graphics, or buttons.

FIGURE 4

Wrap-around Text with One-Half Inch White Space

In 26 words or less, HyperCard is "an innovative data handling environment. HyperCard lets you put information in stacks of index cards, then easily link elements of the data in amazingly complex ways" (Keith Thompson, *InfoWorld*, October 5, 1987). This brief description is accurate, but does not begin to explore the scope of HyperCard's capabilities. For example, the data that you manipulate can be text, graphics, synthesized or digitized speech, or the contents of a CD-ROM or videodisc. You can gain access to these data with the click of a mouse button, and you can move through millions of bytes of information in a matter of seconds. And unlike other traditional databases, HyperCard does not restrict you to linear-sequential organization. You can bound through your data in any way your heart desires by clicking on predetermined text, graphics, or buttons.

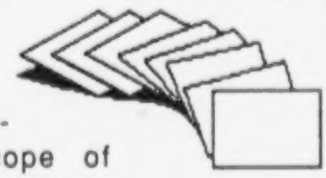


FIGURE 5
Text with No Graphic

In 26 words or less, HyperCard is "an innovative data handling environment. HyperCard lets you put information in stacks of index cards, then easily link elements of the data in amazingly complex ways" (Keith Thompson, *InfoWorld*, October 5, 1987). This brief description is accurate, but does not begin to explore the scope of HyperCard's capabilities. For example, the data that you manipulate can be text, graphics, synthesized or digitized speech, or the contents of a CD-ROM or videodisc. You can gain access to these data with the click of a mouse button, and you can move through millions of bytes of information in a matter of seconds. And unlike other traditional databases, HyperCard does not restrict you to linear-sequential organization. You can bound through your data in any way your heart desires by clicking on predetermined text, graphics, or buttons.

Title:

Educational Computing Social Foundations: A Symposium

Authors:

Nancy Nelson Knupfer
Robert Muffoletto
Marina Stock McIsaac
Alfred Bork
Randall Koetting
Andrew R. J. Yeaman

EDUCATIONAL COMPUTING SOCIAL FOUNDATIONS

Abstract

The purpose of this symposium is to present a collection of six related papers and viewpoints about the social, political, and economic issues surrounding the use of computers in schools. This session will make the audience aware of some of the major social issues surrounding the use of computers in education, cause them to reflect upon the effects of particular research findings and applications currently being implemented in education, and provide direction for those in positions to implement educational innovations.

Each of the six investigations will provide an important perspective on educational issues related to the use of computer technology in schools, its conceptualization, practices, effects, or implications. We hope to provide a collection of ideas that have been traditionally ignored in teacher education and graduate studies, but beg to be addressed. It is from this perspective that we wish to form the contents and context of this symposium.

Each presenter will address a particular aspect of a critical social issue related to the presence of computers in the schools. These issues include historical perspectives on technology and education, teacher education, reskilling of teachers, distance education, equity, change and reform in both pedagogy and institutional structure, concepts of the individual, power and control, machine as expert, issues surrounding courseware evaluation, the effects of computers on teachers, students, curriculum and knowledge, and the question of aesthetics and creativity in educational computing. Papers will be available for the audience.

The papers will be a collection of essays written by experts from universities throughout the United States in both education and educational technology. The symposium will appeal to teacher educators, administrators, teachers, community members involved in education, and students in teacher education, educational computing, or educational technology.

Title:

Educational Computing and Teachers: Changing Roles, Changing Pedagogy

Author:

Nancy Nelson Knupfer

EDUCATIONAL COMPUTING AND TEACHERS:
CHANGING ROLES, CHANGING PEDAGOGY

Dr. Nancy Nelson Knupfer
Arizona State University

Stimulated by rapid technological advancements, the widespread diffusion of microcomputers across all levels of our economy and society has created a demand for substantive computer education. Although parents, government, and school officials recognize that students must prepare for citizenship in a computer-based society, the mere acquisition of computers for the schools is but one small facet of the much larger and more complex task of implementing this potentially significant innovation in education. The careful planning and proper implementation of informed computer education programs, free from unreasonable expectations, will determine the value of this innovation (Goodlad, 1975; Romberg & Price, 1981). Further, the success of any such implementation effort will depend largely upon the teachers who determine the daily school activities (Cuban, 1986). Indeed, the personal and professional relationships among teachers, students, and administrators can have a greater influence on the future of any proposed innovation than the most generous outlays of money and machines.

An accurate view of educational change and implementation of innovations considers the environment within the schools (Berman, 1981; Berman & McLaughlin, 1978; Sarason, 1982). For example, the introduction of computers might change the grouping of students or the amount and quality of interactions between students and teachers, or the very pedagogy of lesson structure. Further, changes in teachers' roles are emerging, which in turn create changes in their relationships with other teachers and with administrators. The current implementation of educational computing is a "social interaction" approach which draws its strength from the practical ideas promoted by teachers and from the directed support of the administration, parents and local community (Knupfer, 1987).

The structure of official authority over computer education varies and can affect implementation. Teachers assume different roles in the change process; some are resisters, while others are active change agents (Knupfer, 1987). This variation of attitudes toward technology can create bridges or gulfs between faculty members, and certainly leads to a hierarchy of power concerning decisions about computer implementation within school buildings, within school districts, and sometimes throughout an entire state.

Implementing computer technology in schools has brought about some obvious changes along with changes which are more subtle, yet carry a full impact. These changes affect teaching pedagogy, the relationship between teachers, the emphasis of inservice training, specific teaching duties, and changes in institutional structure.

Determining Access

Determining access to computers is based upon a number of factors. Some of those factors include the number and location of computers, and the type of software available. Factors which are less obvious are the skill and comfort level of the teacher along at least four dimensions: familiarity of with hardware; familiarity with software; comfort with the different needs of instructional computing as it affect lessons and classroom structure; and, time to prepare.

The number and placement of machines influences access and could possibly scuttle the implementation effort. An insufficient supply of machines predisposes the school in favor of group use at the expense of individualized instruction, contrary to the objectives of the curriculum. In studying the elementary schools, Becker (1985) found that although locating computers in laboratories promoted students' enthusiasm for school, it also fostered the dominance of usage by "above-average" students. On the other hand, placement of computers in the classrooms led to usage by fewer teachers and for a narrower range of applications than the laboratory location. Becker's study also found that placement of computers within libraries promoted more equal usage of computers between "above-" and "below-average" students, but that teachers reported less positive social outcomes and low enthusiasm among students. Placement of computers in each classroom, in a computer lab, or in a central location such as the library without planning and preparation can lead to undesirable results.

In addition to these material factors, existing myths and prejudices about computer use can produce decisions by the school, the teachers, or the district that directly affect equity (equal opportunity for all students) by setting questionable criteria for access. Often, a few highly interested students and teachers monopolize access to the machines by exploiting popular misconceptions about the complexity of computers, thereby constructing psychological or social barriers to entry into their privileged domain. Also, the school's laudable dedication to the special needs of remedial or gifted and talented students can end up restricting access by establishing budget and curriculum priorities that reserve computers for these special groups.

Existing prejudices cover more than just computers, too; for example, ingrained deficiencies in the school's curriculum or the presence of social discrimination, both originating long before the machines arrived, can combine with and even intensify prevailing opinions about computers. Curriculum guidelines that determine competence by grade level alone artificially group students and create unreasonable variations in accessibility.

Certainly, the speed of technological change has intensified problems of equity. The computer's sudden appearance in the classroom, enhanced by its reputation as a highly-touted symbol of the new age in communications, is often misinterpreted as the sign of real, substantive change in education. This false impression encourages both distorted notions about the computer's usefulness and unrealistic expectations about the students' and teachers' ability to learn from the innovation. Schools can get swept up in the computer revolution and lose sight of the machines' potential and limitations, finally losing control of the implementation.

Considering the Curriculum

When developing their computer curricula, schools filter information through criteria shaped very much by familiar institutional patterns and imperatives. For example, curriculum decisions that exclude the entire teaching staff with the significant exceptions of a few well informed teacher experts may suit the current temperament of some staff members, but this approach can stifle discussion about the merits and liabilities of computer education and thereby fragment the implementation, an obvious contradiction between intended integration and the actual practice (Donovan, 1984).

Currently popular opinions about general issues in education have also intertwined with issues in computer education. The public's demand for more teacher training reflects an understandable concern over the efficacy of the schools, but it also leads to hasty decisions and poor planning. As schools rush to respond by requiring more intensive training of teachers in computer operations and by acquiring more software, the original objective of the implementation fades into the background. Nearly a decade ago Sheingold, Kane, Endreweit, and Billings (1981) warned that a sober assessment of the students' needs must temper the school's reaction to public outcry to prevent the sacrifice of the implementation's outcome to only one dimension of the entire process of implementation.

Many programs currently used in educational computing do not fully utilize instructional strategies to best facilitate learning. If this technology is to be truly beneficial and remain a valuable part of our educational system, we must use what history tells us of good teaching techniques and sound innovational strategies, and then apply critical planning and evaluation throughout implementation. It is imperative that computers do not determine curriculum, but that schools formulate their curriculum first and then consider how computers can best serve the instructional objectives and activities of that curriculum.

Critical Nature of Teaching Roles and Pedagogy

The expanding literature on computer education reveals a diverse field of information cluttered with bits of educational theory, curriculum design and implementation plans. However, the best laid plans for implementation of computer curricula depend upon the interaction of the teachers, students, and administrators to make the plans successful. According to Sherif (1970), it is important to realize that the locus of change lies in the interaction between people and that the most effective changes in institutions and attitudes are products of interaction. Sherif's statement emphasizes the importance of realizing the pertinence of social interaction theory and addressing the study of implementation of educational computing from within that context.

Computers and computing are rapidly achieving a prominence in American society, economy, and culture that schools must recognize. Instructional computing challenges the traditional instructional focus, possibly transferring it from the individual teachers to a multitude of instructional sources (programs). This challenge places unusual stresses on the classroom teacher and amplifies the importance of the teacher's attitude toward computers. It is the teacher who lives with the innovation and it is ultimately the teacher who will accept or reject, implement successfully or fail to implement technology in the classroom.

Indisposition to rapid change is a feature of the educational system. Resistance rises from several sources, but these may be classified under the general headings of "reason" and "habit." As with any entrenched institution, the educational system is suspicious of changes that inject new patterns of behavior and threaten traditional beliefs. Computers represent change, but not in a form most educators totally understand. The introduction of computers requires modifications in behavior and instruction that may be unanticipated by the staff; preconceptions about the complexity of computers, the sudden realization that hours of training in a new skill could displace previously important activities, and the zeal of the machine's advocates all can induce fears that this new tool is incompatible with the vital working habits that lubricate personal and professional relationships. Resistance to the change naturally follows. The profession's natural conservatism is only one possible source of resistance, however; the staff's dedication to the institution's objectives, a loyalty based on reason as well as habit, automatically renders suspicious any proposed changes. Although advance preparation for change is no guarantee that resistance will disappear, it will disarm some potential sources of opposition and increase the likelihood of more reasoned argument about an innovation's possible costs and benefits.

The implementation of computer technology is a process of mutual adaptation as the schools blend their individual characteristics with the requirements of a new educational tool. For instance, the incorporation of computers to assist with routine administrative paperwork can produce an interest in training students to operate the machines, which in turn leads to proposals for formal instruction in a number of applications. As students and teachers become familiar with the machines, new demands for new uses arise, and the school is forced to respond.

Previously, computer education was limited to understanding how to use a complicated machine, but technological developments have shifted the limit from the hardware to the human imagination. Imagination about how to use computers in education can lead to creative implementation ideas. Those teachers with creative ideas will continue to stimulate students' interest in the computer, but those teachers who fall behind will soon face the problem of declining student interest as the novelty of computing wears thin.

In general, there is a positive relationship between teachers' expectations and students' achievement. Teachers' attributes influence their thoughts and behaviors toward students. Therefore, teachers' opinions about computer usefulness will influence their motivation to use instructional computing with their students.

Although more computers are being placed in schools, the new equipment has not necessarily improved education. A study completed during the pre-microprocessor age, for instance, warned of the educational system's intrinsic resistance to change and concluded that the rigid, decentralized yet inflexible structure of education stifled any meaningful innovation in pedagogy and curriculum that might have arisen from access to computer (Oettinger, 1969, in Sheingold, et al., 1981, p. 3).

Perhaps the reason that we do not know enough about implementing computers into the schools is because rapid technological changes have outpaced teacher training and curriculum development (Becker, 1986). Changes often seem to take place for no apparent reason, but are based upon innovation for its own sake, with no clear direction. Past experience demonstrates that rushing to place equipment in schools without adequate planning can lead to failure of an innovation (Cuban, 1986; Fullan & Pomfret, 1977; Goodlad & Klein, 1970).

The successful assimilation of computer technology into education must occur along four related dimensions. The first level of assimilation, understanding the process of educational change, must precede the second level, planned implementation of an innovation. But any practical implementation also assumes the third dimension, an understanding of the broad possible applications of computers to education. Finally, the successful integration of computers to education requires its acceptance by a large, diverse, and at times fragmented constituency; therefore the pressures arising from limited availability and uneven distribution of this educational innovation must be overcome in favor of equal access to and quality time at the computers.

Many influential proponents of instructional computing have stressed the issues involved with the process of change, implementation, and equity but have given short thrift to the role of the teachers. Yet, the teacher is essential to the successful implementation of computers, or any school change, into the daily classroom activities (Cuban, 1986). It is the teacher who must adopt computers and then adapt them to curriculum goals and classroom needs (Cuban, 1986; Fullan, 1982).

The teacher's role in schools and school change is so important (Cuban, 1986; Fullan, 1982; Sarason, 1982) that educational change depends on what they do and think (Sarason, 1982). Several factors must be considered in assessing the teacher's role in educational change. The teacher's role must be central to any change; because the teacher lives and works in a classroom with its own built-in imperatives and social culture, the teacher's real working conditions must be taken into account; finally, reformers must realize that the teacher will judge the acceptability of any change by its conformance to current needs and objectives, and not according to some agenda that is foreign to the teacher's experience. In short, teachers must help to guide educational change, and not be its victims.

Powerful implementation of computers into the curricula calls, at most, for a change in the structure of schools and, at least, for a change in the structure of daily lessons. Whether this will happen and how it will happen depends upon the classroom teachers. Even the most traditional teachers who use computers effectively in their classrooms are finding that once they become secure with their ability to implement the technology, the very structure of their teaching changes. Instead of being the source of information, the teacher becomes the facilitator of information, guiding the student in the direction of a solution and perhaps encountering larger issues along the way. In solving the steps of a problem, the students often become more knowledgeable about the topic than the teacher. In order to be successful at this method, the teacher must be willing to release the control of learning to the students, and feel secure in a different role. In such a role there is potential for the students and teacher to learn together, to learn by discovery, to access information beyond the intent of the original lesson plan, and to explore new types of projects.

The balance of incentives and disincentives from the perspective of informed teachers helps explain the results of any change or implementation effort. The teacher's viewpoint of need, clarity, and personal cost/benefit ratio will determine the implementation process. The criteria teachers use to evaluate the change must tip the balance in favor of the change at some point early in the implementation process.

The computer is not just a new piece of equipment added to the classroom, but it calls for changes in the organization of teaching, classroom management, and social interactions during the learning process. Supporting the teacher throughout the implementation process, from planning through evaluation, revision, and ongoing implementation, is critical to the success of instruction computing efforts.

References

Becker, H. J. (1986, June). *Instructional Uses of School Computers. Reports from the 1985 National Survey. Issue Number 1.* Center for Social Organization of Schools, The Johns Hopkins University.

Becker, H. J. (1985, March). *How Schools Use Microcomputers: Summary of the First National Survey.* Center for Social Organization of Schools, The Johns Hopkins University.

Berman, P. (1981). Toward an Implementation Paradigm. In R. Lehming and M. Kane (Eds.), *Improving Schools.* Beverly Hills, CA: Sage.

Berman, P. & McLaughlin, M. (1978). *Federal Programs Supporting Educational Change, Vol. VIII, Implementing and Sustaining Innovations.* Santa Monica, CA: Rand Corporation.

Cuban, L. (1986). *Teachers and Machines: The Classroom Use of Technology Since 1920.* New York: Teachers College Press.

Donovan, B. F. (1984). *Power and Curriculum Implementation: A Case Study of an Innovatory Mathematics Program.* Paper presented at meetings of the American Educational Research Association.

Fullan, M. (1982). *The Meaning of Educational Change.* New York: Teachers College Press.

Fullan, M. & Pomfret, A. (1977). *Research on Curriculum and Instruction Implementation.* Review of Educational Research, 47 (1), 335-397.

Goodlad, J. (1975). *The Dynamics of Educational Change.* Toronto: McGraw-Hill.

Goodlad, J. & Klein, M. F. (1970). *Behind the Classroom Door.* Worthington, OH: Charles A. Jones Publishing Company.

Knupfer, N. N. (1987). *A Survey of Teacher's Perceptions, Opinions, and Attitudes about Instructional Computing: Implications Regarding Equity.* Doctoral Dissertation: University of Wisconsin.

Sarason, S. B. (1982). *The Culture of the School and the Problem of Change.* Boston, MA: Allyn and Bacon, Incorporated.

Sheingold, K., Kane, J., Endreweit, M., & Billings, K. (1981). *Issues Related to the Implementation of Computer Technology in Schools: A Cross-Sectional Study.* New York: Bank Street College of Education.

Title:

Teachers, Computers, and Power

Author:

Robert Muffoletto

TEACHERS, COMPUTERS AND POWER

by

Dr. Robert Muffoletto
The University of Northern Iowa

As teachers implement computer technology into the schools, they are realizing a change in the power structure inherent in daily classroom life. The very nature of their lessons reflect the shift in power from their own authority to the authority of computer software. Expert systems are being developed which do no less than analyze and direct students paths of learning through the software.

Along with teachers who have developed a sense of confidence in their ability to implement computers are those teachers who feel insecure with technology. At either end of the spectrum, teachers must come to grips with their own role in relation to the computer. If the machine is understood as the expert, what role will the teacher have? This paper will critically address questions about knowledge, where it comes from, how students receive it, and how teachers determine their role in the learning process in relation to the computer.

Title:

Economic, Political, and Social Issues that Affect the Growth of Distance Education

Author:

Marina Stock McIsaac

Economic, Political, and Social Issues That Affect the Growth of Distance Education

Marina Stock McIsaac

Distance education, in which the instructor is removed by time and place from the student, offers the promise of serving the needs of the world's growing population without the massive expense of constructing additional classrooms. Such programs rely on technologies which are either in place or are being considered for their cost-effectiveness. Distance education programs are particularly beneficial for the many people who are not financially or physically able to attend traditional classes.

During the past forty years the world population has doubled to over five billion people, most of whom want to be literate and want greater educational opportunities for themselves and their children. The majority of this expanding population is in Asia which faces massive problems of poverty, illiteracy and disease. In most developing countries, such as Bangladesh, distance education offers the promise of a system of information distribution through which new ideas, attitudes and understanding might begin to "ooze through the layers of the disadvantaged environments" (Shah, 1989, p. 9). Drawing upon the well known model of the British Open University, countries such as Pakistan, India, China, and Turkey have combined modern methods of teaching with emerging technologies in order to provide low-cost instruction for basic literacy and job training (McIsaac, Murphy & Demiray, 1988).

Interest in the field of distance learning is accelerating rapidly with the development of new communication technologies. In addition, many economically developing countries have embarked upon distance education programs by adopting existing models and course development strategies from North America, Australia, Great Britain, and Europe. In the past ten years, record numbers of students in developing countries have gained access to higher education through distance education programs (Rumble & Harry, 1982). However, in many cases, developing nations that are attempting to apply established programs from other countries find it difficult to adapt programs that were designed to meet the needs of a different culture.

Distance educators, particularly those in developing countries, are facing the implementation of massive education programs with little research to guide them. Because of fewer resources in these countries to guide in the development of distance education programs, the lack of appropriate research poses a particular problem.

Telecommunication networks now circle the globe, linking people from many nations together in novel and exciting ways. As the borders of our global community continue to shrink, we search for new ways to improve communication by providing greater access to information on an international scale. Emerging communication technologies, and telecommunications in particular, provide highly cost effective solutions to the problems of sharing information and promoting global understanding between people. In today's electronic age, it is predicted that the amount of information produced will increase exponentially every year. Since economic and political power is directly related to access to information, many educators like Takeshi Utsumi, President of GLOSAS (Global Systems Analysis and Simulation) have worked to develop models of the "Global University" and the"Global Lecture Hall" which provide resources allowing less affluent countries to keep up with advances in global research and education.

Lack of education breeds inequality among less affluent people and only serves to widen the gap between the "haves" and "have nots". Access to education has long been recognized as the single most important factor for achieving economic and social equality for citizens of developing countries, minority populations in developed countries, and women seeking equal opportunities. Communication technologies offer these traditionally underserved populations the potential for greater equality by providing them with access to global information. There are economic, political and social issues which have direct impact on how the growth of distance education will occur.

ECONOMIC. The most important consideration for the majority of developing countries is economic independence. It is in many of the economically developing countries that the largest distance learning projects are undertaken. A top educational priority for many such countries is to improve the cost effectiveness of education and to provide training and jobs for the general population. The lack of financial resources available for conducting research, particularly prior to embarking on a massive distance education plan, is a common problem. In many cases investing money in research is perceived to be unnecessary and a drain from areas in which the money is needed. Time is an additional problem, since programs are often mandated with very little start-up time. In the interest of expedience, an existing distance learning program from another country may be used and revised but many times this does not adequately answer the needs of the specific population.

One solution to the lack of adequate research available locally has traditionally been the donation of time and expertise by international organizations to help in developing project goals and objectives. The criticism of this approach is that visiting experts seldom have adequate time to become completely familiar with the economic, social and political factors influencing the success of the project. A second, and more appropriate solution, has been to train local experts to research, design and implement sound distance learning programs based on the needs of the particular economy (Tekin & Demiray, 1989: Shah, 1989). Developing nations prefer to train their own experts and direct their own projects. This preference for locally initiated and conducted research can be seen by examining bibliographies and references.

POLITICAL. One goal of education, particularly in developing countries, is to support the political organization of the country and to develop good citizens. Distance education programs which endorse this priority will have greater chance for success. National political philosophies and priorities are found reflected in the diversity of distance education programs around the world. These programs conform to prevailing political and economic values.

In China, for example, the objectives of distance education are to meet the needs of the developing economy and to assist in the national modernization program, part of a massive catch-up effort (Jianshu, 1988). Thailand, noted for its successful distance education program at Sukhothai

Thammathirat Open University (STOU), responds to the national priority of preserving and developing local culture by producing high quality text books in the Thai language. In Turkey, the directive for Turkish higher education expressed in the Higher Education Act of 1981 was a reform measure legislated as a reaction to the political turmoil and civil disturbances which had shaken the Turkish government in the 1970s. This political situation was reflected on university campuses to the point that they were the scene of fighting and even gun battles. The establishment of the distance education program at Anadolu University is an answer to that political directive (McIsaac & Murphy, 1988). Researchers, whether local or foreign , must be sensitive to a country's political environment if their work is to be seriously considered, implemented and widely distributed.

SOCIAL. Researchers and distance learning administrators need to analyze the social conventions of a country and incorporate those conventions into any distance education program before it can be successful. Such social conventions as extended hospitality, differing perceptions of time and the perceived importance of the project can all affect the credibility of the program and, ultimately, its success. In particular, the social classes and roles of women in society must be considered when making recommendations for distance education programs (McIsaac & Koymen, 1988).

The social, political and economic implications of Distance Education are global. Not only do these issues affect education, they affect the entire fabric and variety of diverse cultures through their use of communication technologies. Only if technology is used to enhance prevalent social, educational and political goals will it fulfill its greatest potential. Once again we have the opportunity to use new technologies to solve some of the worlds pressing equity issues. Paradoxically, distance education offers the promise to provide well designed instruction to help remove the distances between people.

REFERENCES

Jianshu, Z. (1988). The development of higher distance education in China. Developing Distance Education. Oslo: International Council for Distance Education, p.450-452.

McIsaac, M.S. & Koymen, U. (1988). Distance education opportunities for women in Turkey. International Council for Distance Education Bulletin. 17, p.22-27.

McIsaac, M.S., Murphy, K.L., & Demiray, U. (1988). Examining Distance Education in Turkey. Distance Education. 9 (1), p. 106.

McIsaac, M.S. & Murphy, K.L. (1988). Turkey: Does distance education meet national educational priorities? Developing Distance Education. Oslo: International Council for Distance Education, p.307-309

Rumble, G, & Harry, K. (eds.) (1982). The distance teaching universities. London: Croom Helm.

Shah, G. (1989). A distance education strategy for Bangladesh. International Council for Distance Education Bulletin. 19, p.9-14.

Tekin, C. & Demiray, U. (1989). A short profile of the first graduates of the open education faculty, Anadolu University, Turkey. International Council for Distance Education Bulletin. 19, p.65-69.

Title:

The 'History' of Technology and Education

Author:

Alfred Bork

THE "HISTORY" OF TECHNOLOGY AND EDUCATION

Alfred Bork

ABSTRACT

There is effectively a standard history of how computers and related interactive technologies are used in educational environments. This history can be the history of an entire country, or it can be the history of a particular group within a country. While there are variants from country to country, and from group to group, it is standard enough that it bears considering. The history, in many ways, is not encouraging. Many of the earlier stages can be considered regretable. Both schools and training show similarities in computer patterns, but schools receive more attention.

The history of technology in education, world wide, is still an incomplete history. In the last section of this paper I discuss a desirable final stage of this history, if we are to move toward a use of technology that improves education. We have serious problems all over the world in our educational systems. Technology offers us the promise, but not the certainty, of being able to cope with those problems. Technology can enhance learning. The last stage in my history reflects what needs to be done to improve educational systems in the world.

INTRODUCTION

Interative information technologies have been widely discussed and used in learning environments of all types. But the results are often disapointing or poor. A technology with great potential is often poorly used.

So far, there seems to be only a little learning from previous mistakes. The tendency is for a new group or a new country to start with patterns previously used that are now recognized by sophisticated users as incorrect. But perhaps early stages and mistakes are necessary for later developments. And some aspects of early stages are useful for later development. By bringing these ideas and this developmental pattern into the open, it may be possible to avoid constant repetition of mistakes. There is some sign that this can happen.

I do not mean to imply that development involving technology in education is a strictly linear process, although it does seem that some stages precede other stages. As with any complex human activity there is some oversimplification in viewing this complex process linearly. Perhaps a history of this kind might better be written in hypertext! But in much of the description that follows I will talk about the history in a linear fashion, since I am dealing with a linear medium, text. The variation from group to group also confuses the patterns. And statistical data about the progression of steps is not available, and perhaps not really useful.

Many of the features of this history are like those seen with other innovations, but I will not review the more general situation. My attempt is to represent honestly the advantages and disadvantages of interactive technology in all types of learning activities.

The situation is, for convenience, grouped in three stages. First I will discuss beginning stages in the history. Then I will discuss intermediate stages, stages that usually follow the beginning stages. Finally, I will talk about the desirable future of this history.

Again the reader is warned that the outline presented should not be taken in too linear a sense. There are many variations in this history in individual situations. But the patterns are common. The reader might want to compare these stages with a similar history of the use of books in education.

BEGINNING STAGES

The stages discussed in this section represent the most common beginning activities in the use of computers and related technology in education. These stages can occur together, although the drive toward computer literacy, in its simpler forms, may be later than the others examined. We can see this pattern both at the individual school level, and often at the level of a whole country.

1. "Lets Get Lots of Hardware."

The drive to acquire computers, to move as many computers as possible into schools or universities, is typically the first strong movement we see for technology in education. The acquisition of hardware is the major consideration in this stage. Training activities sometime reflect this early stage also. Teachers given a video projector, but no education in video, would not find the equipment useful.

Teachers and administrators are involved in this strong push toward as much equipment as possible, as quickly as possible. For some, this is perceived as a way to gain visibility. But the main forces involved are probably the commercial world, and the parents. The commercial world is involved for obvious reasons, the desire to market their machines. The parents are driven by another consideration, the vague notion that there is something wrong with a school or university that does not "have" computers. Usually no one involved, the school officials, the teachers, the sellers, or the parents have a clear idea of what the computers are to be used for; only equipment counts. The pattern is different if the school or country cannot afford the hardware; in a sense, the cases of this kind are fortunate, because they can avoid the purchase of large amounts of useless, and soon obsolete, hardware.

Sometimes there is "a" use for the computers, a single use that is promoted heavily. This is a slightly more advanced stage but it is still primarily a hardware oriented stage. The philosophy is "let's get many computers all running X, and train the teachers so that marvelous things will happen in our school." The life of the typical student is hardly affected by the new hardware, and this software thrust is often short-lived.

Like many stages, this stage does not die, but continues to repeat on and on. We can still see today either the pure hardware acquisitions, or the hardware acquisition driven by one piece of software. This stage is lacking curriculum and learning consideration, an essential component of educational environments that I will frequently return to in this discussion.

We can think of possible historical analogies to this situation. We might speculate that in the early days of printing schools acquired printing presses, with the notion that they would make considerable difference in how things were done. Schools today may acquire video players for all rooms, with little consideration of what educational tapes are available, how they can be obtained, and the role they play in student learning. Unfortunately, educational issues take second place to technological issues in these considerations. That also happens with computers when the emphasis is primarily on acquiring computers; getting hardware.

Within the United States public primary and secondary schools we now have about two million computers. One could question whether it was wise to buy that many computers, given the limited effective use of those computers within classes.

It is not necessary for everyone to proceed this way. In Japan, where both software and hardware technology exist, we see a conscious effort to avoid putting computers into classes, at least as an early stage. In other countries computers have been avoided too, but primarily because they were too expensive, or difficult to produce, for the countries involved. This last group may be the luckiest!

2. "Lets Teach Languages"

Once the hardware has been acquired in many of these schools there is distinct embarrassment. What is it to be used for? What do students do? The situation is desperate.

One type of software is almost always available, the simple programming languages. They come with the computer in many situations so one common use of these initial machines is to use them for teaching programming. Learning a computer language is usually not what the parents, the teachers, and the administrators have in mind, but it is possible, and like many things in technology, if something is possible, it often happens without consideration of whether it is desirable. So students learn a computer language, or a fragment of a language.

Given that programming is to be taught, and given that teacher and administrators are good at finding reasons for doing something, we begin to see rationales developed for students learning to program. We hear statements like "almost everyone will need to program in the future." Or statements like "one can only understand computers if one writes programs," or "programming increases problem solving capabilities." These are all dubious, vague assertions with little experimental backing, although the third may eventually prove, with good curriculum units, to be true.

Undoubtedly, programming is useful for some people, and perhaps even for everyone, although that is yet to be demonstrated. Many of the advocates of teaching programming relate it to learning general problem solving strategies. But the research on problem solving have not demonstrated transfer of problem-solving capabilities. Problem solving in one area is often unrelated to problem solving in another area, it would appear. Nevertheless, there may still be some interesting possibilities in that direction, to be discovered in the future, as suggested.

My view is that far more interesting learning environments, and material, for programming are needed. We can begin to see some possibilities with the development of the Martino the Robot material in Italy, and the Informatics material developed by the USSR Academy of Sciences and Moscow State University. Both of these projects are curriculum-based, not language-based. More work is needed.

The question of whether programming should or should not be taught is at least an open question, and a question that has many associated subquestions: "how" it should be taught, "to whom" it should be taught, "when" it should be taught, and "what" it should replace. When programming is taught in schools there are three major problems. First there is inadequate curriculum material typically for the teaching of programming. Second, the teachers in schools typically know little about modern programming or software engineering. Third, the languages most likely to be used in teaching programming in schools are poor, low-level languages.

I will not comment on the lack of curriculum material for learning to program, except to say that most of the textbooks in this area are written by people with little previous programming experience. The situation with regard to teachers is more obvious. Modern programming is not a simple activity. The software engineering principles involved are not something that can be learned by beginners in a few weeks of a summer program. We would not dream of certifying people to teach mathematics based on a sole mathematical experience of a few weeks in a summer institute, yet we do a similar thing in programming. A competent teacher of programming needs just as much training as a competent teacher of mathematics. (We are fast approaching the point in the United States where we do not have competent teachers of high school mathematics, too. Recent figures show that only a little more than half of our mathematics teachers are certified to teach mathematics; that is a separate issue.)

The issue of languages requires more attention. There too there is a "history," repeated in many situations. Almost inevitably the initial language is BASIC, although this may be slowly changing. Few computer scientists believe that there is any value in anyone learning BASIC. BASIC was an early language, which did not meet many of the demands of later software engineering. It develops bad habits in the students, almost inevitably in practice, habits extremely difficult to overcome at a later time. The original language has diverged in many directions. There is adequate commentary already, but BASIC continues to be widely taught.

I do not think that the situation is much better with Logo in classes today. Logo does have the possibility of program structure, through its use of procedures, so it is superior in that sense to many variants of BASIC. But Logo is an old language at this point, with a strange syntax, compared to more recent languages. Logo in classes turns out to be about 99 percent turtle graphics, with little to do fundamentally with Logo, since this graphic facility was a later addition after the language was developed. Turtle geometry can be and has been added to almost any language. Most teachers do not understand Logo, and teach accordingly. So results in typical schools are poor.

In the United States, the Advanced Placement exam in computer sciences, leading to college credit, has had a beneficial effect on what language is used, but still the fundamental issues raised above are not touched. The Advanced

Placement exam requires a structured language, Pascal. Since over 95 percent of introductory university computer science courses are based on Pascal, it is not surprising that the Advanced Placement exam in computer science uses Pascal. But the problems mentioned above, with teachers, are still often present. More curriculum material is available.

I remind the reader that I still regard the teaching of computer languages as an open issue. More research is needed, and more experimentation with different ways of learning about programming and problem solving. But even if one ignores all these problems associated with "lets teach languages," this use of computers has little effect within the school environment. Students are studying algebra, physics, English, history, and many other topics. If the computer plays no role in those courses, it is unlikely to have much total effect on the educational system. Hence, the thrust to teach languages, even if it were carried out in an extremely successful fashion, would still represent a limited use of the computer in education, and one using student time that might better be devoted to other purposes. There is general agreement about this in the technical community, but nevertheless languages continue to be taught, often poorly.

3. "Let's Teach Computer Literacy"

This historical development is often a stage beyond the two just discussed. It starts, often, as another name for the teaching of programming languages, one that seems more acceptable to many people. The argument is similar to that for language teaching: that everyone will need to be computer literate in the society of the future, because computers will be widely used in all activities. Parents, teachers and administrators tend to agree on these issues.

But agreement is lost when one needs to consider what the content of a computer literacy course should be. Computer literacy is like motherhood in that most people are in favor of it. But unlike motherhood, it does not have a clear and precise definition. People agree that they like computer literacy only because they do not explore carefully what it is that they have in mind by the term, "computer literacy."

Computer literacy is something of a bridging mechanism from the beginning stages of the history of computers in education to the intermediate stages I discuss next. At one stage computer literacy is indistinguishable from the teaching of programming languages perhaps at a simple level. Indeed some textbooks for computer literacy are fundamentally programming texts. All the considerations in the last section apply.

But, as suggested, computer literacy can have a wide variety of meanings. At the lowest level, it can be simply "can I turn the computer on, and insert a disk, and run a program." It may include the use of business-related tools, such as spread sheets and word processors. The meaning of computer literacy continues to evolve, in all countries.

One aspect of computer literacy does seem important, but difficult to get across to students. People need to consider the ethical and moral problems associated with the use of computers. Computers exert a powerful effect on our society. They can exacerbate major social, personal, and economic problems, or they can contribute to the solution of these problems. But the discussion of these issues is handled, if at all, superficially in most computer literacy courses.

Another frequently stated goal of computer literacy courses is to motivate students. Student motivation, in all areas, is important, as it can increase quality time on task, an important factor in how much students learn. My experience in looking at schools that claim to have a computer literacy program is that a fair amount of this turns into game playing with computers. We get various rationales, primarily the notion that somehow the game is overcoming barriers that the student may have in using computers. I do not see these barriers, particularly with children.

Good user-friendly computer material runs with students with no attempt at trying to convince them that computers are good for them. Or the issues of student excitement are often raised, in using game-like programs. One could equally argue that pinball machines, free candy, or open sex are valuable within classrooms, because they excite the students.

Computer literacy courses have also come to emphasize the learning of common business tools, such as spreadsheets. This is discussed in a later section of this paper.

Toward the end of this historical phase the term computer literacy begins to fall into ill repute. Sometimes it is replaced by other names. In Europe and in many other countries the name Informatics is popular, although it often denotes a higher level course than one associated with the term computer literacy.

4. "Let's Train the Teachers"

It is quickly realized that almost no teachers currently in schools have had any acquaintance with computers. Indeed, the teachers now coming out of schools of education have almost zero acquaintance with computers, because very few schools of education anywhere in the world are in a position to deal with this question adequately. So a movement begins to arise with emphasizes teacher training. As with many of the other stages considered, this is seen as a panecea. We continue to spend large sums of money in this direction, with little in the way of positive results.

This movement is different in different stages of the history. When it comes first in this early stage, it is often concerned with languages and computer literacy. It also follows some of the stages to occur later in this material, such as acquaintance with tools.

The major difficulty with all this is what the teacher should be trained to do. Typically the content of teacher training in technology reflects the stage of the history, as suggested. The fact that there is no coherent use of the computer in education, as will be consistantly pointed out here, raises great difficulties in the issues of preparing teachers to use computers effectively.

To some extent what happens with training teachers about computers now might be thought of in the following fashion. We decide that the overhead projector is to be a major use in education. We run courses on how to use the overhead projector, but we don't supply teachers with much in the way of adequate material, so these courses are content free. We would not expect under these circumstances that courses of this kind would be too valuable.

This suggests that teacher training will not be successful until we have adequate curriculum material using the technology. At that point we will also need something else that is lacking at present, good materials for training the teachers to use these technology-based courses.

NEXT STAGES

At this point the history becomes somewhat more confused, in that not all of these stages are pursued by all groups. Furthermore the order becomes more variable than that just suggested. Often "national policies" of a country, after the initial failures represented in the first stages, begin to come into play, determining these new directions. But the net effects may be similar to grassroots approaches coming from teachers.

This middle stage in our "history," has a variety of possibilities. There is no clear-cut order in these possibilities and arguments erupt about the "correct" order. Various philosophical positions favor one or the other stage, often almost at the level of fads. So although succeeding sections are numbered, these numbers should not be taken to show strict historical order.

1. "Let's Use Advanced Hardware"

The first stage in this history was the drive to acquire computers. Middle history reflects a similar hardware intensity but the grounds have shifted. Now more powerful equipment is considered.

A variant of this hardware drive occurs in prestigious universities. There it is simply not the equipment that is wanted, but the most advanced (and expensive) equipment that the university feels that they can own. Often large amounts of money go into purchasing this equipment, and developing system software associated with it. I won't mention names, but examples will be well known to many readers. Education takes second place; most of the students are unaffected by this advanced hardware, and most classes continue as they were, without the new computers. It is often argued that these 'experiments' help determine future directions, but the experiments seldom are done carefully enough, and are too individualistic, to provide such data.

Another aspect of this rush to use the latest equipment concerns the use of new technology related to the computer, such as the videodisk, nd the various variants of compact disc. Some groups rush from one equipment to another, looking the the pot at the end of the rainbow.

This is not to say that these devices will eventually not prove to be useful in learning. I believe they will. But if we always proceed by jumping on the bandwagon of the latest device, the computer will never have any major effect on education! We never seem to use the technology that we have, but rather many individuals are eager to proceed to a newer technology. We *already* have the technology to make major improvements in education in all countries and at all levels. This will be further considered in the final stage of this history.

Again, we see a mistaken emphasis on the hardware, and a lack of attention of what can be done in courses. The computer, with all the new peripherals, still should be considered the major device, because it alone brings interactivity and individualization. But the question of what technology should be used should not be driven by the latest piece of equipment around. Rather, it should be driven by pedagogical considerations. Decisions made without taking into account the full learning context are likely not to be adequate decisions. Thus, if we examine interactive video, the combination of the computer and the video disk, there are some areas where visual information will be extremely important, either for direct content or for motivational reasons, and there are other areas where it may be of little consequence. Deciding in advance that I am going to use equipment X is often a poor decision.

2. "Let's Develop Small Programs for Use in Standard Courses"

The reaction in this case is to a comment already made in the earlier section. The computer, used in the vein suggested in the early history, has no effect on most of the curriculum, and little effect on most students. Most of the courses taught are in subject matter areas. If the computer is to affect education significantly it must be used in these subject matter areas. Educators slowly begin to realize this.

Hence, we see the beginning of development of computer based learning material. Normally this early development is at a cottage industry level. A teacher, knowing a little bit about programming, writes a program for his or her students. Large numbers of teachers proceed in this way. The programs are all small, because the teachers have limited time and capability in this area. Then organizations are founded to take the teachers' programs, modify them, improve them, and make them available, perhaps commercially. Thus organizations may begin to develop their own small programs. Many, but not all, of the companies distributing computer based material have histories such as this.

a. Characteristics of Small Programs.

The hallmark of these materials is that they are *small*, not taking students much time. That is, they do not represent any extensive amount of material, but are bits and pieces, minute programs, perhaps occupying an hour or so of student time.

Because of this it is not too surprising that these sequents typically have little effect on the courses in which they are used. A traditional course, based primarily on lectures and textbooks, has a few pieces of computer based material introduced at certain points. Not too surprisingly the overall course is little altered. The student time spent with the computer is likely to be small compared to the time spent listening to lectures, or reading books. So the courses are only slightly altered. Only a small percentage of student time goes to using computer material.

The courses that use such small programs already exist, and are typically based on print material developed many years ago. So their form and content reflects pre-computer strategies. Hence they cannot make full use of the new technologies.

Another problem with many of these small programs is that of quality control. Since they are often developed on an individual scale, without adequate resources, and often by individuals with little experience in developing learning materials, it is not too surprising that the quality is often poor. Yet many of these programs do get used, for the reasons already suggested: the computers are in the schools, and there is some drive to show that the computers are useful in some way to justify the cost of the hardware. Many become commercially available. Some good programs do exist, but too few to improve education for most students, and the programs themselves occupy too little student time to have much effect.

Another interesting, but educationally unfortunate, phenomenon arises in connection with these small programs. Programs that have almost nothing to do with educational content, as it is understood in the traditional fashion, or by those who look toward future systems, become widely used, because they exist, and are perhaps entertaining to children. Thus the enthusiasm for creating computer-based posters, greeting cards, and banners, having little to do with the learning process, reflects the existence of software for doing this. In a sense the software situation is very similar to that with regard to languages, as previously discussed. We have it, so we use it! Since the software exists, it is used, sometimes heavily. Hence the computer has an unfortunate effect on education; it takes away valuable student time from important components of learning, and invests it in material which has little intrinsic learning value.

b. Types of Small Programs.

Several types of these small programs are developed, and again large philosophical discussions break forth about which type is "correct." These philosophical discussions are seldom based on pedagogical or empirical considerations about the needs of students in learning in a course. They are simply widely held sets of beliefs. Nevertheless they are often argued vigorously. What is lacking is much empirical information to support the arguments, in any direction.

- One common type of material is tutorial, units that try to help the student in the learning process.

- Higher level tutorial material attempts to work somewhat the way an individual tutor might work. At its best, seldom realized, this material can individualize the learning experience for each student, and can make learning an active experience. It also plays a major role in maintaining student interest. Groups of excellent teachers can be effective in designing such units. This is the approach we follow at Irvine.

  But little tutorial material of this kind exists at present, and it suffers from the same problems as any small amount of material. So the potentials for this stage in history are greater, but still remain potential. This topic will receive further discussion in the next major section.

  At the lower end of tutorial is drill and practice, giving students practice in using the material. Drill and practice is often derided, yet in many courses, as they are taught today, drill in learning the materials is almost essential. For example, if one looks at almost any calculus course taught anywhere in the world today, these courses require large amount of practice, drill, in the arts of differentiation and integration. Whether this is a reasonable procedure or not deserves further consideration, but it is unfair to criticize drill and practice on the computer when it is heavily used in noncomputer components of the curriculum. Drill and practice with the computer has distinct advantages over the usual drill and practice, such as the possibility of immediate feedback and assistance to students.

- Another type of small program that has its strong philosophical adherents is the simulation. These go under various names. At Irvine we call them "controllable worlds". If they are done in Logo they tend to be called "microworlds". Sometimes they are called "free environments". The notion of a simulation is straightforward. The computer simulates some part of a real or imaginary world, creating conditions that allow the user to "play" with that world, change it, and observe what happens.

  There is a strong pedagogical purpose for simulation, although not often understood. Simulations are most useful in the important task of helping students to build their intuition about particular areas, to build insight. This experiential phase is an important component of learning, often neglected by our traditional processes. Insight can be stimulated by a rich collection of student experiences provided by simulations.

  But the typical small simulations suffer frequently from a fatal flaw. They are not interactive, so they cannot react to what the student is doing or not doing, and so cannot determine whether the student is building intuition, or gaining any other form of knowledge in the process. Student difficulties are ignored in many simulations.

  So a few students may be learning with the simulation, but many students are not. These later students tend to be bored by the material, since they do not see what is happening, or why they are involved in it. The designers of simulations often assume that student help will come from teachers, but given typical class size this expectation is not at all reasonable in most situations. Student motivation is often a serious problem. Simulations are often far more interesting to the developers than to the students! Students do not understand why they are using the simulation.

  I call a simulation that completely ignores what the user is doing, a naked simulation, because it does not have the "clothes" that it needs to be a full pedagogical unit. An effective simulation needs to watch carefully what the student knows and does not know, offering help and assistance. But few of these programs do. Naked simulations are seldom useful with typical students, although they may work for very good students with very good teachers. They need to be clothed in pedagogy. A friend tells me that the term 'naked video' has a similar meaning.

- There is a recent alternative to simulations that also is connected with building student intuition, through providing a wide range of experience. This alternative is the micro-computer based laboratory. The notion is to attach a probe directly to the computer, a probe that measures some aspect of the outside world, such as temperature, distance of an object, etc. As with other types of material, little empirical information is available as yet to how this approach compares with the approach of the "pure" simulation. Many scientists favor actual laboratory equipment, even though it is, as in the case of a micro-computer based laboratory, removed from direct manipulation by the student. Typically these programs produce graphs of whatever the probe measures, or whatever can be calculated from this, versus time.

  As with the simulations, most of the microcomputer-based laboratory programs do not pay attention to what the student is doing. So the problems mentioned for simulations also occur. The assumption once more is that the teachers will provide this additional information, without regard for the fact that in the large typical classroom it is usually impossible for the teacher to do this for everyone.

- Some tutorials and simulations get the word "intelligent" attached to them; but these programs seldom actually use the methods of artificial intelligence. We do find a few experimental programs that do use these methods, and we find other programs without artificial intelligence methods that clothe the simulation. But these are rare at this current stage of the history. The potential, however, is again large.

c. Elitist Programs

A serious problem with many tutorials and simulations is that they only work with a few very good students. The typical student is quickly in trouble. Learning material of this kind deserves to be called *elitist software.*

Given privileged students, or extremely competent teachers, the things that are missing in the naked simulations and simple tutorials can be supplied, at least to a few students. Even here however, the chance of being able to supply the missing details, to everyone in the class, is likely to be small. Too many students are present. Hence, software of this kind often is glowingly reported based on its use with a few bright students and teachers, but fails in typical courses.

3. "Let's Use Authoring Systems"

The drive toward producing small pieces of learning material quickly becomes frustrated, because the teachers involved know little or nothing about programming. Hence, entrepreneurs see a new direction, authoring languages and systems.

Teachers are told that they do not need to know anything about programming. "We have this easy way of allowing you to write learning material, through the magic of computers, where no programming is needed." Anyone can do it, the promoters claim. Languages may start within a professional organization (IBM and Coursewriter, for example).

Authoring languages and systems are heavily promoted and advertised products. At the 1988 ADCIS meeting in Philadelphia, for example, about half of the exhibitors were promoting authoring systems. Many of these are 15 or so years old. Some reappear under different names. It seems that profits are to be made in this area, given the considerable commercial interests that market authoring systems and languages.

Little quality material is written in these systems. If the systems are easy to use, as advertised, then they produce only simplistic material. As they get more complicated they become less easy to use, but provide a general range of material. When this happens, the systems typically are removed from the hands of the authors, and a new level of professionals uses the system. The original advertising promises are then forgotten.

But these authoring facilities are seldom designed by competent computer scientists who know what a complex programming language or developmental environment should be like, and so they are typically not adequate to the task. Further, they often have a restricted built-in view of the nature of the learning process, or they can produce only certain types of materials. They do, however, absorb great quantities of resources, money and time, and so their net effect on the promotion of computers in education has been negative; they use up money that could be better spent elsewhere. The net result of most authoring systems is to make money for the promoters (sometimes).

But almost every novice getting into the field believes, unfortunately, that the first choice is to pick an authoring system or language! New systems keep appearing, suggesting that some systems make a profit! But little effective material is produced with these systems. It is often not realized that authoring systems, as opposed to sets of tools, are not needed to create high quality interactive learning material. The problem, not discussed further, is primarily a specialized problem in software engineering.

4. "Let's Catalogue Existing Software"

Since many of these small low quality programs are developed, the numbers suggest that we need lists of them. So we find organizations that prepare these lists, and distribute them. Sometimes these organizations are called clearing houses; they also may make available some of the material on these lists. Not too surprising, clearing houses list material of variable quality. If anything has potential, it may be picked up commercially, and so will not be a candidate for the clearing house. So these lists, even relegated to particular fields, seldom turn out to be of much use. A few exceptions exist.

5. "Let's Evaluate the Small Programs"

A step beyond cataloguing is some attempt at evaluating the material. This is certainly desirable, because it begins to give some ideas about which small programs might conceivably be useful in which classes. This might be related, in schools, to the state and local frameworks that guide the curriculum.

But evaluation can have many meanings. What is usually implied at this stage of the history by evaluation is peer evaluation, the program being examined by another teacher who did not write it, but who is familiar with the students in the area. The critical issue, however, is how effective the programs are in helping students to learn. Peer evaluations emphasize teaching, not learning.

Many of the materials glowingly described in peer evaluations turn out to have severe problems when ordinary students are involved in a summative evaluation. Materials liked by teachers, as already suggested, may not work with students; this is not unique to computers. While some useful things can come from peer evaluation these evaluations miss the major direction for evaluation, testing with typical target audiences. Summative evaluation, well done, is expensive and time consuming in most situations.

The net result of these evaluations, even though the quality of the effort is poor, is to confirm what has already been pointed out, that there is little effective learning material available in the small program category. The nature of the programs, that they are small, almost assures that this will be the case. The material is not extensive enough to make a real difference.

6. "Let's Teach Students About Tools"

An entirely different approach is seen, often toward the end of this middle stage of our history. The argument resembles what we have already seen in computer literacy, at least in its earlier stages. It goes something like this. Powerful computer tools are now generally available in our society. Students may almost certainly use these tools after they graduate from school or university. Hence, let's be sure to teach them all about the tools that are currently available. As with learning languages, this philosophy can be questioned.

The first thing to note is that the tools usually being referred to, word processors, spreadsheets, data base manipulators, graphics programs, have almost all come from the business community. That is, *none of them* were developed for use in the learning process. I sometimes refer to this stage in the history as "picking up the crumbs of the business table."

a. Business Tools in Classes

Because these tools were developed for business considerations to meet business needs, it would seem almost obvious that they are not ideal for educational purposes. But this seems to have escaped most of the people who pursue this direction. Often the tools used are not exactly those that were developed for business. The reason for this is that the computers they are running on are computers that businesses refuse to buy! Thus weaker forms of the business programs, imitations for the weaker computers, are often developed for the school or university environment. Schools often do not use the "best" business tools, even when they have the appropriate computers.

Where should the tools be taught? Often the computer literacy course is revived in this context, and it becomes a course for learning about word processors and other tools. But we also see use of tools within subject-matter classes, and some of these approaches are useful, at least slightly. Thus a spreadsheet may get used within a mathematics class, or within a class in the sciences. The published literature contains many examples of this kind, and we find enthusiasts for this use.

Perhaps the most important use of such tools within a class is the use of word processing within the learning of writing. Word processors will soon be almost universally used in almost any type of writing activity, so the presence of word processors in courses in learning to write, or that involve substantial writing, is not surprising.

Within these courses it may be the student's responsibility to learn how to use the word processor sufficient to the needs of the course. This is often the case, for example, in junior and senior level English courses in universities, where it becomes almost impossible to survive in the course, given the amount of writing needed, without a word processor. Learning to use the tools is not simple, and often little help is provided.

Certainly the word processor is very useful in writing, and in learning to write. The ability to make corrections easily, without the laborious processes involved in rewriting and retyping entire papers, shows the potential for improving writing. Spelling checkers also play an important role.

The word processor is only one of many computer-based tools that can be useful in the writing environment. Furthermore the learning of writing involves a whole methodology of which word processing is only a small segment. Learning to write has undergone a revolution recently, with emphsis on viewing the entire process of writing, not just the actual writing activity. The notion that there are tools that can assist with the prewriting process, and tools that can help with the post-writing or revision process, are central to using computers effectively in a process approach to improving writing. Yet these tools are often fragmentary, and seldom put together in a full scale curriculum designed for learning writing and for later productive writing. Although there are some exceptions to this, these exceptions are not the most commonly used tools in writing.

b. New Tools

The use of tools does not end with the business tools. New tools are particularly developed for areas of the curriculum.

To some extent these tools overlap types of software considered previously, particularly simulations. As with the business tools, problems arise from the issues of curriculum.

c. Curriculum Context of Tools

The emphasis on a curriculum context, just expressed in the use of word processing within the teaching of writing, is an important and neglected consideration for most of these tools. Having the tools alone, whether they are word processors, spreadsheets or other things, is seldom sufficient.

What is needed is a whole collection of software to use with the tools. This software needs to address the following questions.

- How is the student to learn to use the tools? How many hours of valuable class time in a writing course are going to be devoted to teaching people the mechanism of various tools? Or is there some more effective way of doing this?
- How do the tools integrate with each other, and how do the students understand that the important thing is not the tools themselves, but the curriculum material that one is trying to get across?
- How do the tools fit with other aspects of the curriculum.
- Does the tool make any attempt to determine if the student is learning anything?
- Does the tool offer any individualized assistance to the student who is in difficulty? Or does it assume that the already overworked teacher will work individually with each student in a large class?

All these questions are seldom answered, because the tools are plucked from business applications, with little pedagogical context surrounding them. The problem is not too unsimilar to the problem of the naked simulation, mentioned above. The focus is not on learning, and on helping students to learn, but on the technology. A friend likes to say that any piece of learning without a curriculum context is like a "one night stand." The analogy seems to me a good one; no full rich relationship is developed.

But even if we can answer all these questions carefully, and develop a full curriculum context for the use of tools, the question remains as to computer use in many other areas of the curriculum where the tools may not be as vital, but where other kinds of technology-based learning material may play a more important role. The notion that somehow education needs to use only the tools that have been developed for business is a very peculiar notion, reflecting the poverty of education today.

7. "Let's Use Networks"

One stage that is eventually reached in many countries and in many individual projects is the stage that focuses on the use of networks, at first local networks, and then long distance networks. That is, the standard personal computer is no longer seen as adequate, but the student also needs connections between computers, and the resources that networks can provide. It is not necessary that every user be networked all the time; many long distance networks assume that only occasionally will there be a connection made by phone, sometimes without knowledge of the student.

Initially the use of networks becomes almost purely a matter of convenience and bookkeeping. The convenience comes because it is no longer necessary to deal with all the floppy disks. The bookkeeping comes because now it is possible to keep better records about what the student has or has not done. The next stage is the use of electronic mail and bulletin boards, often poorly coupled with the curriculum; these activities can be valuable if integrated with the curriculum.

But seldom has any learning material been developed that *demands* the network, other than the use of electronic mail and bulletin boards. A few exceptions exist, such as the National Geographic/TERC science units. Although there would seem to be some interesting possibilities for pedagogical developments that could take place in network environments, and could not take place in other environments, these possibilities have scarcely been touched. So networks become an expensive addition, not contributing much to the learning situation. But like many of the early stages in the history, they reflect a potential that might be realized if the final stage of history is ever attained.

8. "Let's Develop Management Systems"

Sometime, usually early in the second stage, the notion of using the technology to provide course management, with or without any of the rest of the material in the course involving the computer, is broached.

The argument for management systems is reasonable. Teachers spend much time with paperwork, time they do not consider interesting. If the management system can indeed cut down this time, then it has considerable appeal for the teachers.

At the simplest level the management systems become simply the replacement for the teacher's record book, a way of recording the grades and progress of the course by the teacher directly. Thus the teacher gives tests, and instead of entering grades in a written form in a record book, they are entered online, hopefully with a friendly query-type program. Perhaps the computer does some mathematics, such as averaging the grades. Such use is of minimal interest.

More advanced forms of management systems may tie in directly with other material on the computer, so that the information is recorded automatically. Systems of this kind do exist. Probably we can expect more of them in the final stage of the history. The integrated system available from such companies such as Computer Curriculum Corporation and WICAT combine management systems with extensive curriculum material.

The purpose of management systems is not given much consideration. Often it is only the teacher that is thought of in recording the information, and supplying various reports. But equally, and perhaps of greater importance eventually, is a management system that supplies help directly to the student. Such a system is necessary if the student is to understand what he or she needs to work on further. Management systems can thus offer a variety of interfaces, for different users.

FINAL STAGE OF THE HISTORY

Up to this point there has been I believe historical accuracy in telling what has happened. However, like any archetypical history that tries to encompass an entire movement, the individual details may be different than those described. The account we have so far brings us roughly up to the present. The computer, the videodisk, and the compact disc are still having little impact on improving education. Some of the reasons for this have been discussed in the early sections of this paper.

But this history is incomplete; we have not reached the desirable final stage for effective use of computers within the classroom environment. The next stage of the history has not yet taken place, in any country or in any organization, although some are closer to it than others. In this section, I give a quick view of what I regard as the next stage of the history. Many of my papers and books expand this in more detail. The reader is referred to those for expanded details.

This last stage in history might be described as "let's develop full curriculum using computers." It uses all the previous experience, including some historical forerunners to the last stage that I have not mentioned.

a. Visions of Future Educational Systems

I find surprisingly little discussion of the desirable forms of future learning systems, of any type. This discussion is much needed; it is difficult to know how to improve schools if we do not have a target for the future. Most of the use of technology in education has ignored this issue.

b. New Courses and Curricula

The only way interactive information technologies can be used effectively in education, just as the only way the book could be used effectively, is to develop entirely new courses. These new courses need to assume from the beginning of the development process that *all* educational media, including those from modern interactive technology, are available. Such material would lead to new curricula. It would go far beyond the bits and pieces of technology-based material now available.

These courses could well differ greatly in form, content, and teacher role from existing courses, because they have been developed later. Content needs change because of our changing society. New forms for course structure are suggested by new learning technologies. Exciting possibilites exist that would have been impossible with the older learning media.

But little in the way of resources have been devoted to full curriculum development based on interactive technology. Indeed most of the money associated with computers in education, in the United States, is still invested in the earlier stages already mentioned in the first two parts of this paper, even today. Some full courses have been developed: the logic and set theory courses at Stanford University, the physics course at the University of California, Irvine, and the Writing to Read course for young children. But these courses are a drop in a very large bucket, if we consider all the systems of education in existence.

This lack of full-scale interactive courses is curious. Hundreds of full video-based courses have been developed, some slightly interactive, but only a very small number of computer-based courses. We cannot realize the full potential of interactive learning technology without newly designed courses. I find it hard to understand why video has had so much more support, given that most video-based courses are not interactive.

c. Teacher Training

The fact that we are considering courses that are quite different than previous courses, including different teacher roles, indicates the great importance of educating the teacher. Current efforts at educating the teacher in something called "technology" are typically unsuccessful, because the technology is not tied down to pedagogical situations. We do not train teachers to use books!

The training of teachers must be associated with each of the new courses developed, and the development of this material for teachers must be considered an integral part of the development of the courses. One of my friends, Jaques Hebenstreit, has emphasized this recently in private correspondence. He comments, "The constructive use of computers is not a problem in hardware and not even a problem in software. It is a problem of making teachers confident."

d. Evaluation

So the focus of the final stage will be on full course development, so that we can rebuild schools and universities with technology-based courses that were not possible with the older technologies. One can see glimmerings beginning in a few countries, but still it is a stage that we are perhaps not ready for at present.

These full courses also need to be accompanied by full summative evaluations. The importance of testing learning material with students has already been stressed in discussing evaluation in the previous stage. It is desirable that several such courses will be available in each area, so that these courses can be compared for student effectiveness not only with the conventional courses, but with each other. Writing to Read has been evaluated in this sense, but such evaluations are seldom done.

e. Implementation

The use of these new courses in all areas of learning is a complex task. Educational institutions of all types are conservative, and do not change easily. But we know, from industry experiences, something about implementation of major changes in education. That experience will be useful here.

f. Conclusions

The potential for this last stage in our history is exciting. We have great problems in education, and it is only with this redevelopment of our courses and schools that we can begin to face these issues and proceed through to this final stage.

Title:

Educational Technology Curriculum Theory: Toward a New Language of Possibility

Author:

J. Randall Koetting

EDUCATIONAL TECHNOLOGY CURRICULUM THEORY AND SOCIAL FOUNDATIONS: TOWARD A NEW LANGUAGE OF POSSIBILITY

by

Dr. J. Randall Koetting
Oklahoma State University (Temporarily at Reno, Nevada)

There is an increasing body of critical literature in the social foundations and curriculum theory that could identify a new language of possibility within the applied field of study of educational technology. This literature could move toward the reconceptualization of taken-for-granted assumptions and provide new directions for research and study in our field. It also has the possibility of broadening the dialogue of those in our field with diverse educators within other areas of study. This paper will explore current literature in these areas of study (social foundations, curriculum theory, and critical pedagogy) and will suggest the importance of reconceptualizing educational technology within a framework of a new language of possibility. The piece will be presented from a conceptual, theoretical perspective.

Title:

Visual Discrimination, Learning, and Page Layout: Wrap-around, Run-around, and Transparent Text

Authors:

Nancy Nelson Knupfer
Marina Stock McIsaac

Visual Discrimination, Learning, and Page Layout: Wrap-Around, Run-Around, and Transparent Text

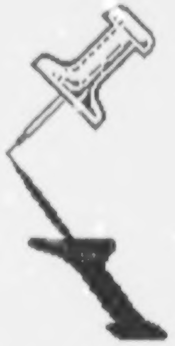
As desktop publishing revolutionizes printing and publication, sophisticated computer programs offer the novice a wide range of techniques formerly available only to professional layout artists. Now anyone can design and produce their own instructional materials by transforming existing art into professional quality work that can be pasted into any variety of text formats.

Several popular software programs are noted for versatility and ease of layout design. They not only feature run-around text, but now boast wrap-around text and transparent text features as well. Run-around text is that which is "boxed" around the graphic, maintaining a square border around the visual (see Figure 1). Wrap-around text is that which flows smoothly around the graphic and "hugs it tight", leaving an irregularly shaped border (See Figure 2). Transparent text is that which has a graphic laid over it; the reader reads the text through the graphic (See Figure 3). These formatting features are described by software publishers as desirable software advancements, which allow the designer to quickly and rapidly communicate ideas to the reader. Examples of the use of run-around, wrap-around, and transparent text can be found in consumer magazines, newspapers, and instructional materials. In many instances, legibility appears severely restricted by the haphazard mixture of text and graphics.

Desktop publishing computer programs have individualized the page layout process and have provided sophisticated graphic enhancements for computer users who have no training in either instructional development, page layout, or graphic design. People who use such computer software to design instructional materials are faced with a vast array of newly developed features with no guidelines for their effective use. Guidelines are sometimes provided but they are often confusing. In one guide to basic design, a designer is told to "Frame your pictures with white space...White space separates headlines from surrounding body copy and artwork, making headlines easier to read" (Parker, 1987, pp 42-43). The examples show both run-around and wrap-around features being used.

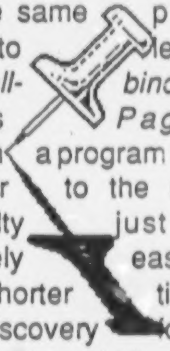
Teachers and trainers who use computer programs to design instructional materials and integrate visuals in the text are asking for guidelines. How do these new desktop publishing enhancements affect the legibility of text? What is the optimal placement of text and graphics? How and where should white space be located to enhance reading comprehension? At what point does the integration of the text and visual

FIGURE 1
Run-Around Text



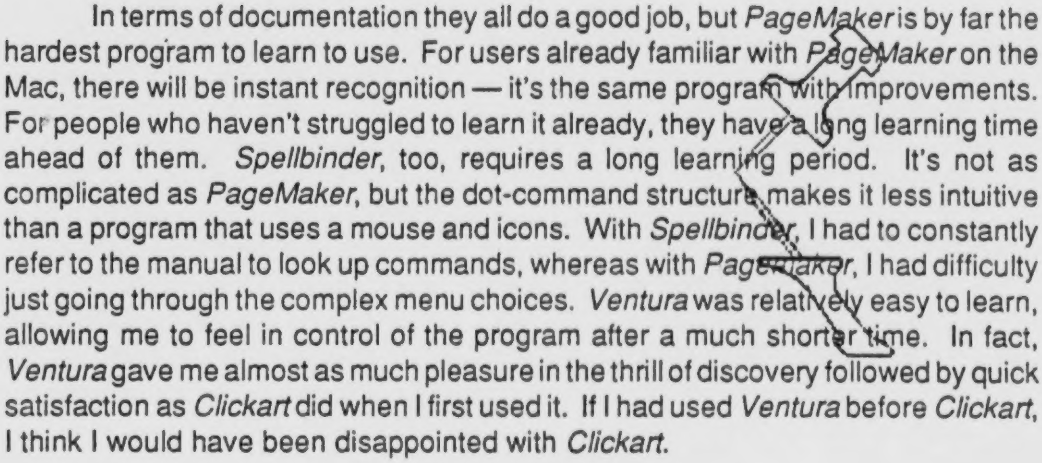
In terms of documentation they all do a good job, but *PageMaker* is by far the hardest program to learn to use. For users already familiar with *PageMaker* on the Mac, there will be instant recognition — it's the same program with improvements. For people who haven't struggled to learn it already, they have a long learning time ahead of them. *Spellbinder*, too, requires a long learning period. It's not as complicated as *PageMaker*, but the dot-command structure makes it less intuitive than a program that uses a mouse and icons. With *Spellbinder*, I had to constantly refer to the manual to look up commands, whereas with *Pagemaker*, I had difficulty just going through the complex menu choices. *Ventura* was relatively easy to learn, allowing me to feel in control of the program after a much shorter time. In fact, *Ventura* gave me almost as much pleasure in the thrill of discovery followed by quick satisfaction as *Clickart* did when I first used it. If I had used *Ventura* before *Clickart*, I think I would have been disappointed with *Clickart*.

FIGURE 2
Wrap-Around Text



In terms of documentation they all do a good job, but *PageMaker* is by far the hardest program to learn to use. For users already familiar with *PageMaker* on the Mac, there will be instant recognition — it's the same program with improvements. For people who haven't struggled to learn it already, they have a long learning time ahead of them. *Spellbinder*, too, requires a long learning period. It's not as complicated as *PageMaker*, but the dot-command structure makes it less intuitive than a program that uses a mouse and icons. With *Spellbinder*, I had to constantly refer to the manual to look up commands, whereas with *Pagemaker*, I had difficulty just going through the complex menu choices. *Ventura* was relatively easy to learn, allowing me to feel in control of the program after a much shorter time. In fact, *Ventura* gave me almost as much pleasure in the thrill of discovery followed by quick satisfaction as *Clickart* did when I first used it. If I had used *Ventura* before *Clickart*, I think I would have been disappointed with *Clickart*.

FIGURE 3
Transparent Text



In terms of documentation they all do a good job, but *PageMaker* is by far the hardest program to learn to use. For users already familiar with *PageMaker* on the Mac, there will be instant recognition — it's the same program with improvements. For people who haven't struggled to learn it already, they have a long learning time ahead of them. *Spellbinder*, too, requires a long learning period. It's not as complicated as *PageMaker*, but the dot-command structure makes it less intuitive than a program that uses a mouse and icons. With *Spellbinder*, I had to constantly refer to the manual to look up commands, whereas with *PageMaker*, I had difficulty just going through the complex menu choices. *Ventura* was relatively easy to learn, allowing me to feel in control of the program after a much shorter time. In fact, *Ventura* gave me almost as much pleasure in the thrill of discovery followed by quick satisfaction as *Clickart* did when I first used it. If I had used *Ventura* before *Clickart*, I think I would have been disappointed with *Clickart*.

---

adversely affect legibility? How close to the visual should the text be placed?

PURPOSE OF THE STUDY

The purpose of this study was to determine the effects of three electronic text variables, run-around, wrap-around, and transparent text, on reading speed and comprehension. Results mights suggest some guidelines for the placement of graphics within electronic text.

Many researchers have investigated layout, typographic cueing, and presentation of graphic materials in instructional text. For example, Hartley (1985) determined that page size affects other layout decisions. Recent studies comparing the effects of justified and unjustified text are inconclusive. For example, Jonassen (1982) reported that in some circumstances, right justified text inhibited comprehension when compared to unjustified text. Yet, Campbell et al. (1981) and Muncer et al. (1986) determined that reading speed and comprehension were not significantly influenced by justification versus nonjustification. Jandreau et al. (1986) investigated the effects of varying spaces between phrases in sentences. And, others have examined certain design features common to printed and electronic text in the presentation of graphic materials (Alesandrini, 1987; Poggenpohl,1985; Winn,1987; Winn & Holliday, 1982). Graphics designers emphasize the importance of using white space for balancing text, positioning graphics as the dominant visual element, and improving the aesthetics of the page layout (Parker, 1987). And, in 1987, Hartley, Isaacs, and others noted a need for published research on the effect of various text-wrap enhancements with either printed or electronic text. There is little reported research on graphic placement in instructional text and none that has produced guidelines for instructional use of graphics with the emerging computer software programs.

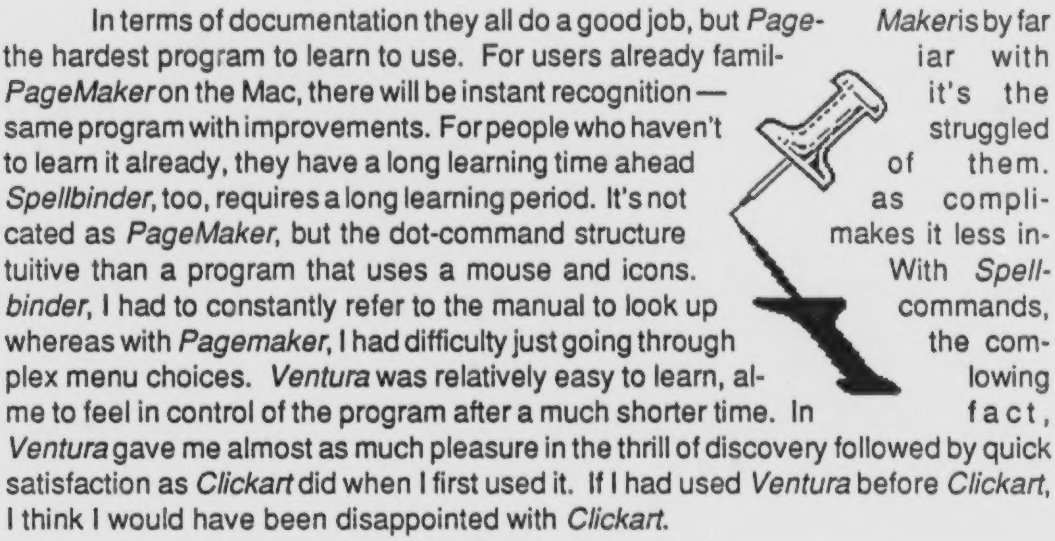
Although certain features of electronic text can be transferred to printed text, there are many differences between the two which suggest that separate studies be conducted. The size of the screen or page, the type size, and the ability to scroll are but a few of the differences which affect print and electronic text comparison studies. Because of these differences, a programmatic series of investigations is needed to examine the effects of these recent technological enhancements in each of the two media formats.

Knupfer and McIsaac (1988) responded to that need by studying the effect of run-around and wrap-around text, within one and two column formats, on reading speed and comprehension. That study found that the number of columns had no significant effect, however the text wrap did make a significant difference. Reading speed was faster and comprehension was better with wrap-around text, as shown in Figure 2, than with run-around text, as shown in Figure 1. In a follow-up study, Knupfer and McIsaac (1989) tried to determine if that difference was due to the amount of white space between the graphic and the text, or due to the shape of the white space surrounding the graphic. The follow-up study tested the variable of white space between the graphic and text by comparing two sets of reading materials which both used wrap-around text only; the difference was that one wrapped the text tightly against the graphic, as shown in Figure 2, while the other conformed the text to the shape of the graphic but left approximately one quarter inch of white space between the graphic and the text, as shown in Figure 4. There were no significant differences between wrap-around text with no white space and wrap-around text with one quarter inch of white space.

This study continues the investigation of text-wrap variables by adding an additional text-wrap feature, transparent text. If the results prove similar, guidelines should then be established to assist teachers and trainers in the appropriate use of this new type of desktop publishing feature. It is hypothesized that the previous results will be replicated and further, that the transparent text will show greater interference with reading speed and comprehension.

Recent work by Morrison, Ross and O'Dell (1988) examined attributes of computer based instruction, and attempted to affect certain attributes by regulating text density for print and screen display lessons. Results suggested that print had more advantages than the computer

FIGURE 4
Wrap-Around Text With One Quarter Inch White Space



In terms of documentation they all do a good job, but *PageMaker* is by far the hardest program to learn to use. For users already familiar with *PageMaker* on the Mac, there will be instant recognition — it's the same program with improvements. For people who haven't struggled to learn it already, they have a long learning time ahead of them. *Spellbinder*, too, requires a long learning period. It's not as complicated as *PageMaker*, but the dot-command structure makes it less intuitive than a program that uses a mouse and icons. With *Spellbinder*, I had to constantly refer to the manual to look up commands, whereas with *Pagemaker*, I had difficulty just going through the complex menu choices. *Ventura* was relatively easy to learn, allowing me to feel in control of the program after a much shorter time. In fact, *Ventura* gave me almost as much pleasure in the thrill of discovery followed by quick satisfaction as *Clickart* did when I first used it. If I had used *Ventura* before *Clickart*, I think I would have been disappointed with *Clickart*.

---

screen in achievement, completion time, and learning efficiency for a statistics lesson. Therefore, the present study is limited to the examination of print-based, computer-generated text .

METHOD

Subjects.

One hundred, thirty-two undergraduate students from an introductory computer literacy course at a major university in the Southwest, participated in the study. The elective course was a 400 level course which was offered to meet a General Studies numeracy requirement. Students were sophomores, juniors, and seniors who had little or no previous exposure to computers. The study was conducted at the beginning of the semester and all students in the course participated. Students had little or no prior knowledge about the topic presented in the study materials.

Materials.

Readings. Readings included approximately six pages of informative text about hypercard and desktop publishing. These materials were chosen because the topic was relevant to the subjects and because the reading level was geared for the general public as opposed to technical readers. The topic was suited to a novice audience with no prior knowledge about either hypercard or desktop publishing. The text was analyzed using the Flesch-Kincaid Readability formula and was found suitable for college students at the 11th and 13th grade levels.

Three sets of reading materials were used for this study. All included the same text and graphics arranged in single column format, but each set utilized a different text wrap respecting the graphic. The three sets included one with straight-framed graphics (see Figure 1, Run-Around Text), one with text tightly fitted around the graphics (see Figure 2, Wrap-Around Text), and one in which the graphic was placed directly over the text (see Figure 3, Transparent Text).

Although the graphics placement within these three sets of reading materials might not follow all of the elements of good design, it is important to know that the way in which they were developed is indicative of what practitioners will design. A sixth grade teacher who had a masters degree in the design and utilization of educational media, and who helped to instruct a desktop publishing course, was very much involved in the design of these reading materials. Therefore, any flaws in the materials design are potential problems that other teachers and trainers might exhibit as well.

Pretest and Posttest. To measure the students' comprehension of the reading material, a pretest and a posttest were developed from key concepts which the authors extracted from the readings. No test items were dependent upon the information contained in the graphics. The tests were developed and validated during a previous phase of the project, and were found to be reliable at the .76 level. The pretest and posttest were identical with one exception; the posttest contained an additional six-question attitude section. The attitude section included questions regarding the students' judgement of the difficulty of content, the degree to which graphics interfered with their comprehension and reading speed, their degree of interest in the subject, and their overall understanding of the material after reading it. These attitude questions were measured on a Likert-type scale and were not included in the posttest comprehension score.

Procedure.

All subjects were given a pretest to determine their knowledge of the subject prior to reading. There was no time limit and students were free to take the time they needed to complete the test. Next, the three sets of reading materials were randomly distributed to the students so that each student received one set of articles. Students were instructed to read through the text one time without hurrying. When subjects were finished reading, each of their reading times was recorded. Immediately following the reading, the subjects were given the posttest for comprehension. They were allowed as much time as they needed to complete the test.

RESULTS AND CONCLUSIONS

The design of the study was an analysis of variance. The independent variable was the text-graphic format (run-around, wrap-around, or transparent text) and the dependent measures included reading speed and gain in comprehension as measured by the difference between the pretest and posttest scores. The alpha level for significance was set at .05.

Table 1 shows the mean scores for the pretest, posttest, gain, and time elapsed for each of the treatment groups. Although the ANOVA showed no significant difference in either comprehension or speed among the three treatment groups, the table does show some patterns that may be worth noting. First, it appears that the wrap-around treatment group had a higher mean gain than the transparent treatment group, which in turn had a higher mean gain than the run-around treatment group. Second, it appears that the wrap-around treatment group also had the fastest reading speed, followed by the run-around group and then the transparent group. We can conclude that the wrap-around treatment group had the best comprehension and the fastest reading time, even though those results are not outstanding enough to be statistically significant.

TABLE 1
Mean Scores

	Pretest	Posttest	Gain	Time
Run-Around	13.85	45.47	31.62	13.74
Wrap-Around	13.55	51.91	38.36	12.59
Transparent	12.57	47.76	35.19	14.51

We might consider the possibility of interference in the scores of the transparent text group. Although the three text-wrap forms were randomly distributed, the transparent group started off with a mean pretest score that was a full point lower than the other two groups. There is no explanation for this occurrance. One might speculate that the transparent group took more time looking at the graphics than the other groups, although one student commented in a post study discussion that she was unaware of any graphics being there as she read, suggesting that she read right through the graphics. We venture to guess that the lower comprehension scores on for the run-around text group might be because the text is more difficult for the eye to follow due one of three possible reasons: 1) there are larger gaps of white space that the eye must jump across to continue reading, 2) there are variations in the amount of white space that the eye must jump across to continue reading, or 3) the text justification might interfere with reading continuity.

Irregular lines of text around a graphic and graphics placed over text did not confuse the readers in this study. The new enhancements currently available in desktop publishing software, particularly the wrap-around features, were not found to inhibit reading.

Because a significant difference in both reading speed and comprehension was found between the run-around and wrap-around treatment groups in a prior study, we suggest replication of this study. Further, we suggest that subjects in each of the three treatment groups be quizzed about their recollection of the graphics to see if they actually look at the graphics as they read or if they concentrate on the text alone. Because graphics are used for different reasons, such as decoration versus information, it could be important to know whether the subjects actually look at the decorative graphics. It is possible that subjects' reading speed and comprehension with the three text-wrap configurations could vary dependent upon the type of graphic used.

Future studies should consider perceptual principles such as proximity, similarity, continuity, and closure. These four processes, by which the mind organizes meaning, depend on how physically close the objects are, how similar they are, whether there is a continuous line to guide the eye, and whether the minimal amount of information is present that is necessary to obtain meaning or closure. Reading speed and comprehension are directly affected by the way the mind organizes meaning from the placement of graphics and text.

REFERENCES

Alesandrini, K.L. (1987). Computer graphics in learning and instruction. In H.A. Houghton and D.M. Willows (Eds.), The psychology of illustration: Vol 2: Instructional issues. New York: Springer Verlag.

Campbell, A.J., Marchetti,, F.M. & Mewhort, D.J.K. (1981). Reading speed and text production: A note on right-justification techniques. Ergonomics. 24, 611-622.

Hartley, J. (1987). Designing electronic text: The role of print-based research. Educational Communication and Technology, 35 (1), 3-17.

Hartley, J. (1985). Designing instructional text. (2nd ed.). New York: Nichols.

Isaacs, J. (1987). Text screen design for computer-assisted learning. British Journal of Educational Technology, 18 (1), 41-51.

Jandreau, S.M., Muncer, S.J., & Bever, T.G. (1986). Improving the readability of text with automatic phrase-sensitive formatting. British Journal of Educational Technology. 17 (2), 128-133.

Jonassen, D.H. (1982). The technology of text. Englewood Cliffs NJ: Educational Technology Publications.

Knupfer, N. N. & McIsaac, M. S. (1989). White Space, Graphics and Text. Proceedings of the 21st Annual Conference of the International Visual Literacy Association, held in Scottsdale, AZ. Oct. 25-29, 1989.

Knupfer, N. N. & McIsaac, M. S. (1988). Graphics and Text Placement: The Effect of Wrap-Around and Framed Text on Reading Speed and Comprehension. Proceedings of the 20th Annual Conference of the International Visual Literacy Association, held in Blacksburg, VA. Oct. 19-22.

Morrison, G.R., Ross, S.M., & O'Dell, J.K. (1988). Text density level as a design variable in instructional displays. Educational Communication and Technology. 36 (1), 103-115.

Muncer, S.J., Gorman, B.S., Gorman, S. & Bibel, D. (1986). Right is wrong: An examination of the effect of right justification on reading. British Journal of Educational Technology, 17,(1), 5-10.

Parker, R.C. (1987). The Aldus guide to basic design. Seattle: Aldus Corporation.

Poggenpohl, S. (Ed.) (1985). Graphic design: Computer graphics [Special issue]. Visible Language, XIX,(2), 173-306.

Winn, W. (1987). Using charts, graphs and diagrams in educational materials. In D. M. Willows and H.A. Houghton (Eds.) The psychology of illustration: Vol.1. Basic Research. New York: Springer Verlag.

Winn, W.D. & Holliday, W.G. (1982). Design principles for diagrams and charts,In D.H. Jonassen (Ed.), The technology of text. Englewood Cliffs, NJ: Educational Technology Publications.

Title:

Theory Building and Educational Technology: Foundations for Reconceptualization

Authors:

J. Randall Koetting
Alan Januszewski

THEORY BUILDING AND EDUCATIONAL TECHNOLOGY: FOUNDATIONS FOR RECONCEPTUALIZATION

There are words, phrases, ideas that can be referred to as contested/problematic concepts, i.e. these words are open to debate and discussion as to their meaning. Not only do the meanings of these words change depending on their use, but they can mean different things to different people. This is not a radical position to espouse until it is given some specificity. For example, the words democracy, capitalism, and politics can be considered contested/problematic, depending on who utters them, or how the utterances are made. In the field of education words like education, curriculum, research, and methods can be seen as contested/problematic. The notions of contested and problematic suggest debate, re-interpretation. They suggest some notion of necessitating dialogue with others as to the establishment of meaning.

The literature within the field of educational technology does not often address itself to the issue of the contested/problematic concept. We tend not to want to think of our field as a battleground for philosophical/theoretical debate. Yet there are contested/problematic concepts in this field. Words such as science, technology, instruction, theory, design, research, etc., can all be seen as contested concepts that are frequently used in the field of educational tecnnology. While there are on-going debates in educational literature in general, there is not much work that looks at differing conceptualizations and frameworks within educational technology. However, when the Association for Educational Communications and Technology in the United States published the Definition of Educational Technology (1977), the intent was not to shut off such debate but rather to promote it.

The limited discussion in the recent literature of the field of educational technology is concerned with the importance of theory to the field (cf. Reigeluth, 1983). The discussion seems to center around three questions: Why use theory? What is theory? and What kind of theories are there? An explicit discussion of one of these questions will lead inevitably to an implicit discussion of the others. In this paper we will discuss the notion of theory as one example of the contested/problematic concept. We will suggest that through the diversity of understandings of this concept, new language can be used to talk about the work of educational technologists. This new language, in turn, will open new possibilities for dialogue and praxis within the field of educational technology.

WHY USE THEORY?

A useful way to enter into a discussion of theory would be to identify the differing ways for arriving at knowledge, or ways in which we know the world. This will help us to understand the nature of theoretical work. This is suggested by the work of Habermas (1971), and thus has a certain philosophical positioning in critical theory. Distinctions are made between the empirical-analytic sciences, the historical-hermeneutic sciences, and the critical sciences. Habermas' notion of interest will also provide a conceptual framework for our discussion.

Habermas' theory of knowledge has three forms, or processes, of inquiry. Knowledge can be arrived at through empirical-analytic science, historical-hermeneutical science, and critical science. These forms/viewpoints of knowledge result in three categories of possible knowledge:

> Information that expands our power of technical control; interpretations that make possible the orientation of action within common traditions; and analyses that free consciousness from its dependence on hypostatized powers. These viewpoints originate in the interest structure of a species that is linked in its roots to definite means of social organization: work, language, and power (Habermas, 1971, p.313).

These categories of possible knowledge thus establish the "specific viewpoints" from which we can know reality in any way whatsoever: "orientation toward technical control, toward mutual understanding in the conduct of life, and toward emancipation from seemingly 'natural' constraints" (Habermas, 1971, p.311). These modes of inquiry with constitutive interests delineate the way in which individuals generate knowledge.

The notion of knowledge constitutive interests linked with the differing sciences situates theorizing within the realm of political discourse. The notion of knowledge constitutive interests means that all our attempts at explaining/understanding our world (context) are non-neutral activities. Inherent in our attempts to make sense out of our world are particular interests.

The idea of interests is not a pure abstraction, nor is it an empirical proposition. Interests are not causes or determinates in and of themselves. They are general orientations/cognitive strategies toward inquiry.

Knowledge constitutive interests are formed in the "medium of work, language and power" (Habermas, 1971, p.313). This means that language is political. The language that we express ourselves in is political. We exercise power and influence through our use of words. But, we can also exercise power and influence by determining the meaning that words have, for ourselves and for others. Thus knowledge is not neutral. The way we go about knowing is not neutral. The way we go about knowing the world happens through a dialectical relationship that we have with the world: we are shaped by our world and we help to shape that world. This also suggests that theorizing about that world is part of a social process, and therefore, theory itself can be considered a social construction. Theorizing, as a social construction/social process, arises out of humankind's desire to explain and/or understand and/or change the human condition.

Theory can serve to help explain both the conscious and unconscious, as they relate to nature, society, and our personal existence. Theory can be seen as a hypothetical proposition which can be proven/disproven through empirical testing. It suggests causal relationships (cause/effect). This notion of "explaining things" is the basis for rational thought. Nomological knowledge is the result of such endeavours.

Nomological suggests lawlike propositions based on the results of the testing of the hypotheses. This form of theory (empirical/analytic theory) has an interest in prediction and control.

Theory can serve to help humankind better understand "the world", and hence ourselves and others within a given context, the meanings attached to social customs, etc. When engaged in this form of understanding/research, the way in which we relate to what is being investigated is seen through different lenses. We use an interpretive mode of understanding/theorizing. This theoretical stance to the world sees reality as a social construction. This form of theory (historical-hermeneutic theory) has an interest in better understanding that social construction through consensual agreement.

We can use theory to gain insight into seemingly "given" realities, and through a process of reflective critique, we can examine the social construction of reality and seek ways to analyze the contradictions found in reality (the "is" and the "ought"). Through a shared vision we can begin to set about the enormous difficulty of changing individual and group context. Hence a different understanding of reality is needed in order to act within that context, to effect change. This form of theorizing (critical theory) has an interest in emancipation. By emancipation is meant the possibility of individuals freeing themselves from "law-like rules and patterns of action in 'nature' and history so that they can reflect and act on the dialectical process of creating and recreating themselves and their institutions" (Apple, 1975, p.126). In this sense, emancipation is a continual process of the "critique of everyday life".

Fundamental differences separate these modes of knowing and theorizing. The differences are of a philosophical nature. Each mode of inquiry has its own understanding of the nature of reality (ontology), the nature of knowledge (epistemology), the nature of questions asked about reality, i.e. what is of value (axiology). The need for theory further arises when there are multiple understandings for the same reality. How do we separate out the theoretical stance that makes sense within our own experience?

WHAT IS THEORY?

Acknowledging the need for theory also involves some understanding of what theory is. Thus far we have spoken in broad and encompassing terms. In this section we want to discuss the word theory in the context of the field of educational technology. Acknowledging the need for theory also involves some understanding of what theory is. The AECT in The Definition of Educational Technology (1977) seems to have endorsed a specific notion of the meaning of the word theory. It is stated that "the term 'theory', while often used colloquially as an antonym for the terms 'practice' or 'practical' has a precise meaning:

1. a general principle, supported by considerable data, proposed as an explanation of a phenomena; a statement of the relations

believed to prevail in a comprehensive body of facts (English & English, 1958, p.551);

2. a principal or set of principles that explain a number of related facts and predict new outcomes based on these facts (Wheeler, et al., 1975, p.638)". (AECT, 1977, pp.20-21).

The AECT's idea of theory is based on the view of the "hard" sciences, or the empirical sciences. The explicit emphasis is placed on the 'facts' upon which some principles are based. A further emphasis is placed on the ideas of explanation and prediction. There is no mention of the concept of understanding. If it was considered at all it seems that it was intended to be implicitly bound to the ideas of explanation and prediction (control). But understanding a particular phenomenon is different than trying to explain or predict it. Understanding something has a different goal than explanation/prediction. Understanding looks toward the whole phenomenon, not the parts. Understanding identifies relations among parts, the interconnectedness, the gestalt. Referring to our earlier discussion, empirical-analytic theory has an interest in technical control. Understanding has a practical interest in consensual agreement, consensus as to the meaning of particular phenomenon.

Since meaning resides in individuals and not in the phenomenon, one must engage in a dialogic endeavour to reach agreement regarding the meaning of something. Through dialogue, agreement can be reached (consensus). This theoretical stance can lead to new understandings of reality, seeing something in a new way. This use of theory suggests multiple interpretations of the same event. This form of theorizing moves back and forth between theory and practice. Theory informs practice and vice versa. This understanding of theory does not necessarily lead to action/change.

Another understanding of theory, i.e. critical theory, has an interest in emancipation. Within critical theory, emancipation suggests change, social and/or personal. Thus theory can be identified by its methodological interests: technical control, consensual agreement (understanding), and emancipation. In other words, theory can be identified by what it is "supposed to do": create nomological knowledge, i.e. law-like propositions that explain/predict/ control; offer a possible understanding of a particular phenomenon (consensus as to the meaning of something); and/or offer the possibility for emancipation.

KINDS OF THEORY

Within the field of educational technology, like other areas of study, theories can be classified as either normative or descriptive. Normative theory has to do with "the articulation and justification of a set of values" (Eisner, 1985, p.49) that identifies a coherent set of beliefs that become the grounds for action within a particular world-view. Normative theory is not concerned with prediction and control, nor empirical verification. An example of such normative theory would be the articulation of a "philosophy of life". As Eisner states, "the roots of such theory begin as humans speculate on what is good in life and worth achieving" (Eisner, 1985, p.50).

Descriptive theory refers to "those statements or concepts that attempt to explain, usually through their power to predict, the events of the world" (Eisner, 1985, p.52). Descriptive theory is a call for action. The emphasis is placed on attaining some goal. Learning theory is an example of descriptive theory, i.e., through learning theory, we identify the processes of learning in such a way that through the application of that theory we can control the learning process, and through that control, we can predict the outcomes of learning.

These conceptions of theory, normative and descriptive, are useful categorizations but they do no close the question "What kind of theories are there?" One way to explore the question "what kind of theories are there?" is to do an analysis of how the word theory is used in various areas of academic study and then look for examples in and draw parallels to its use in the field of educational technology. According to the philosopher/scientist Ernst Nagel (1969, quoted in Kliebard, 1977, pp.262-263), there are four kinds of theory:

1) The law-like theory of the "hard" sciences such as physics and chemistry. This relates most directly to "empirical findings".

2) Theory that is supported by statistical evidence. Theories that have been developed in the study of bio-genetics are examples of this kind of theory.

3) Theory that is an attempt to identify factor variables which are major influences in a field of study. Many economic theories are of this variety. These theories attempt to identify the variables (unemployment rate, deficit spending, etc.) that may impact on the study of economics.

4) Theory as a systematic analysis of a set of related concepts. Therories from sociology that analyze how a given society operates are examples of this use of the word theory.

Nagel's discussion of theory is rooted in the empirical-analytic tradition. His first three uses of the word theory are descriptive in nature and are concerned with explanation/prediction/control. Explanations lend themselves to suggested actions. In his fourth use of the word, theory is reduced to the role of an heuristic, or operational definition, i.e. as a device or tool used to clarify areas for investigation, but not investigation itself. Thus understandings that are arrived at from this perspective are not theories, but can lead to theoretical understanding, but only after empirical investigation.

It is Nagel's fourth sense of theory that we see as promising for our discussion. Theory becomes both a conceptual analysis of a given word and a normative statement of its use. In this sense of the word theory, empirical verification is all but impossible. Since the foundational and hence major questions of education are normative in nature (i.e. "they involve choices among competing value options"), what becomes "critically important is conceptual clarification" (Kliebard, 1977, p.263). This fourth use of the word theory is not reduced to merely heuristics, but becomes the grounding of new understanding and the potential for change (action). Theory is thus not reduced to explanation/control, but has the potential for broadening understanding (interpretation) and expanding the options for action.

CONCLUDING COMMENTS

The political implications of the AECT'S sanctioned "theory" are evident. From our discussion above regarding the historical-hermeneutic and the critical sciences, we believe the empirical-analytic view, posited as the only legitimate way of generating knowledge and hence theoretical positioning, is inadequate for dealing with the complexities of a field of study. The AECT's use/endorsement of this notion of theory has had the effect of closing the door on other uses and discussion of the term theory. The AECT also seems to equate understanding with prediction and control (explanation). This effectively reduces the intellectual investigations that it sanctions to those that can predict. Any notion of emancipation or the hermeneutic sense of understanding is not seen as legitimate. Thus the field can be seen as not only limiting what can be investigated, but it limits the possibilities of learning of those whom research in the field supposedly helps, namely the "clients" (children, the business community, etc.). This results in a prediction/control cycle: professionals are limited to certain types of investigations (those that fit into the empirical-analytic model), and hence, "learners" are subjected to the results of this limited approach.

If we stay where we are in terms of our limited views of theory and theory building, we will cut ourselves off from the cultural sciences and the differing ways in which the world can be interpreted. Again, it is Nagel's fourth use of theory that we see as providing possibilities for our field. The prediction/control cycle can be broken through hermeneutical and critical inquiry. The hermeneutic interest in understanding and the critical interest in emancipation begin to give a radically different language and conceptual basis from which to work. This new language and interest will present new possibilities for dialogue and praxis within the field of educational technology.

Engaging in theoretical debate within our own field can lead to further explorations. These explorations will not be confined to academic matters, but will become part of personal experience. We are concerned with philosophical questions everyday: What is real? How do we know? What is of value? These are questions that effect what we do professionally. They also affect us on a deeply personal level, and become the basis for our praxis.

BIBLIOGRAPHY

AECT. (1977) **Definition of Educational Technology.** Association for Educational Communications and Technology, Washington, D.C.

Apple, M.W. (1975) Scientific interests and the nature of educational institutions. In Pinar, W. (Ed.) **Curriculum Theorizing: the Reconceptualists.** McCutchan Publishing Corporation, Berkeley, California.

Eisner, E.W. (1985) **The Educational Imagination: On the Design and Evaluation of School Programs** (2nd ed.). MacMillan Publishing Company, New York.

Habermas, J. (1971) **Knowledge and Human Interests.** Beacon Press, Boston.

Kliebard, H.M. (1977) Curriculum theory: give me a 'for instance'. **Curriculum Inquiry**, 6:4, 257-269.

Reigeluth, C.M. (1983) Instructional theories in action: lessons illustrating selected theories and models. In Reigeluth, C.M. (Ed.) **Instructional Design Theories and Models: An Overview of their Current Status.** Lawrence Erlbaum Associates, Hillsdale, N.J.

THE NOTION OF THEORY AND EDUCATIONAL TECHNOLOGY: FOUNDATIONS FOR UNDERSTANDING

J. Randall Koetting is currently a Visiting Associate Professor, University of Nevada, Reno. He is on a year's leave from Oklahoma State University, Stillwater, Oklahoma, where he has been a faculty member for eleven years, teaching in the graduate program of study, Curriculum/Supervision. He received his PhD from the University of Wisconsin-Madison, from the department of Curriculum and Instruction. He has a Master's degree from St. Louis University, St. Louis, Missouri, in the area of Educational Foundations, and a bachelor's degree in philosophy from LaSalette Seminary College, Ipswich, MA. His research interests include critical theory, curriculum theory and pedagogy; the social foundations of education; issues that fall under the general heading of multicultural education; and the foundations of technology and schooling.

Alan Januszewski is currently completing his doctoral dissertation in the area of Instructional Design, Development, and Evaluation in the School of Education at Syracuse University in Syracuse, New York. He has been awarded a Master's of Science degree in Instructional Design, Development, and Evaluation from Syracuse University, and a Master's of Science degree in the Cultural Foundations of Education from the University of Wisconsin-Milwaukee. He has taught graduate courses in education at Potsdam College of the State University of New York and Syracuse University. His areas of scholarly interests include the social impact of the use of technology in education; the history and philosophy of education; and curriculum theory and development.

Title:

The Effects of Feedback and Second Try in Computer-assisted Instruction for Rule-learning Task

Authors:

Doris Lee
Patricia Smith
Wilhelmina Savenye

Abstract

The effects of three types of feedback: knowledge of the correctness of response (KOR), knowledge of the correct answer (KCR) and explanatory feedback in combination with or without a second try on error correction in a computer-assisted instructional lesson were investigated, using a 2 by 3 factorial design. The results indicated that explanatory feedback was significantly superior over KOR feedback in improving immediate posttest achievement and also significantly better than KCR feedback in reducing the probability of repeating a practice error on the delayed posttest. The "second try" factor had no significant impact on performance. One of the possible reasons for the insignificant results might be that learners did not maintain a high level of learning motivation due to the difficulty of the instruction. Future research is needed to determine what would be the sufficient amount of feedback information and appropriate difficulty level of the instruction that would maximize learning.

The Effects of Feedback and Second Try in Computer-Assisted Instruction for a Rule-Learning Task

In the past decade educators have emphasized the importance of making instruction more sensitive and responsive to the individual learner. Educators often argue that instruction should be developed in a manner which provides the availability of as much time and resources as required by the learner to attain the learning goal. Computer-assisted instruction (CAI) has been considered as a valuable individualized instructional systems to meet these time and resource demands. The key reason for this belief is because CAI can be designed to individualize the instructional process in a way which can adapt the instruction to each student's individual and changing differences in terms of needs and abilities. One of the ways that CAI can accommodate the individual needs of the learner is to provide feedback to learners as to the adequacy of their performance.

Function of Feedback

Feedback can be employed to communicate with the learner about the adequacy of his performance and, therefore, is crucial for learning. Unfortunately, its mechanism in learning has not been unequivocally determined. For a long time, behavioristically oriented theorists (Skinner, 1961, 1968; Throndike, 1931) tended to assign a reinforcing mechanism to feedback. In contrast, recently, a number of psychologists have pointed out that feedback should be informative and corrective and has a more facilitating effect as error correction than as a reinforcement of correct response (Ausubel 1968; Kulhavy, 1977). Furthermore, Gagne (1987) asserted that feedback functions as a mechanism for supporting effective storage of infomation into memory.

Various studies have investigated whether it is the reinforcing or informational characteristics of feedback that facilitates learning. The research results have repeatedly confirmed that the functioning of feedback in learning is informational, rather than reinforcing (Buss & Buss, 1956; More and Smith 1965; Sassenrath, 1975).

Types of Feedback and Their Effects

Many researchers prefer to treat feedback as a unitary instructional variable, while some indicate that feedback should be classified into various categories based on its instructional purposes. For instance, Kulhavy (1977) pointed that feedback can be regarded as a "continuum" which can range from the simplest confirmational format to the presentation of remedial information or even new instructional materials.

A number of researchers further believe that the different natures and amounts of feedback information have differential influence on learning (Carter, 1984; Gilman, 1969; Roberts & Parks, 1984; Roper,1977). As Carter (1984) reported, feedback is usually constructed into four forms including (a) knowledge of response (KOR) feedback, which contains simplest message such as "correct", or "incorrect" only, (b) knowledge of the correct answer (KCR) feedback, which informs the learner as to what the correct answer is, (c) corrective feedback, which states why an incorrect response is erroneous and why a correct response is right and (d) various embellishments of these three basic types.

Several studies have been conducted to investigate the effectiveness of different types of feedback in order to determine how much information and what kind of information should be contained in feedback. Gilman (1969) examined the effects of five types of feedback (no feedback, KOR, KCR, corrective and a combination of KOR, KCR and corrective feedback) in a CAI lesson dealing with commonly misunderstood science concepts. The results revealed that the most detailed feedback information yielded the best results (i.e., the number of correct responses in posttest). Also revealed in Gilman's study was that KCR, corrective or a combination of KOR, KCR and corrective feedback produced better effects upon immediate retention than did no feedback or KOR feedback. Research results which favored feedback with increased information have also been conducted by Roper (1977) and Roberts and Park (1984). In Roper's study, the effects of no feedback, KOR, and KOR feedback plus the correct answer stated in the context of the question were investigated. Roper found that KOR feedback coupled with the correct answer resulted in the best achievement on the posttest and no feedback was the least effective feedback condition. Roberts and Park (1984) studied the interaction between different amount of feedback information and learners' cognitive styles. In their studies,learners were classified as field dependent or field independent and received a CAI lesson along with either KOR or explanatory feedback. They found that learners receiving explanatory feedback performed better on both the immediate posttest and the delayed retention test than those who received KOR feedback, regardless of their cognitive styles.However, with respect to the effect of feedback with additional information, there are other studies that have failed to yield consistent results. For example, the study by Steinberg (1981) did not reveal significant differences among no feedback, KOR feedback and KOR feedback plus additional explanation. Further, there is little consensus on precisely how much information should be contained in feedback so as to promote learning. Phye (1979) tested the prediction that the greatest amount of feedback information would produce the best immediate and delayed retention effects. Unexpectedly, he found that the minimal feedback information brought the greatest retention improvement on the posttest. According to Phye, a possible explanation for such a contradictory result may be that too much or too detailed information seems to confuse learners.

Number of Tries

In a CAI lesson, second try opportunity is occasionally provided after an incorrect response for the purpose of supplying learners an opportunity in which they can correct an previous error and further employ their corrected mental structure to determine the correct answer. Unfortunately, however, only a few studies have investigated whether the provision of a second try opportunity upon an missed question facilitates learning. For instance, Hodes (1984) investigated the effects of corrective feedback, which informed an incorrect response as wrong, gave an corrective opportunity on the missed question and provided supportive encouragement to try the missed question again, and noncorrective feedback, which only informed an incorrect response as wrong. The results showed that corrective feedback and noncorrective feedback did not yield differential effects upon retention.

The Purpose of Study

This study was concerned mainly with how different types of feedback including

KOR feedback with a simple confirmational message, KCR feedback with the correct answer and explanatory feedback with the corrective information tied with one or two number of tries worked in correcting errors. Primary questions addressed in this study were: (a) What are the comparative benefits of KOR, KCR and explanatory feedback following an error during practice on posttest performance? (b) Does a second try as opposed to a single try provided following an error during practice promote error correction on posttests? and (c) Is there an interaction effect between level of feedback and number of tries?

Method

Participants

One hundred and sixty undergraduate preservice teachers from a southwestern university participated in this study. Among them, 90 percent were female. A pretest was administered to all participants as to (a) determine their prior knowledge of BASIC, which was the content taught in the experiment, and (b) use pretest scores to match students during assignment for treatments in order to reduce the possibility of initial group differences. Participants were grouped into high, medium, and low prior knowledge groups according to their performance on the pretest, creating stratefied random assignment of students to groups. Participants in each of these three groups then were randomly assigned to one of the six treatments. The results from the pretest showed that more than ninety percent of the participants were considered as low prior knowledge students. Consequently, in each of the six treatments, there were only two to three participants who were considered as high or medium prior knowledge students.

With regard to attrition, one student did not take the immediate posttest and twenty- eight students did not return for the delayed posttest. An examination of immediate posttest scores did not yield any indication that attrition was biased by any particular treatment. One error occurred during recording data on to the diskettes. As a result, only one hundred and thirty students successfully finished the experiment.

Independent Variables

Types of feedback and number of tries were the two independent variables involved in this study.

Types of feedback. Specifically, KOR feedback conveyed the simplest type of feedback information to learners. It merely informed the learner whether his response was correct or not. For example, KOR feedback after a correct response was "You are right" and KOR feedback after an incorrect response was "You are wrong". KCR feedback stated the correct answer regardless of the correctness or incorrectness of the response. For example, KCR feedback given after an incorrect response was "The correct answer is ... ". Explanatory feedback included KOR feedback to inform whether a response was correct or not and told the learner why an incorrect response was erroneous. In the case of an incorrect response, the learner could be informed of an error and realized why he made such an error through studying the content of the explanatory feedback. In this study, explanatory feedback never stated what the correct answer was. Primarily, it was used to provide the corrective information that enabled the learner to correct himself or served as a form of remedial instruction .

One and two tries. A single try was provided when the practice question was answered correctly. In contrast, a second try opportunity was given following an incorrect response. It was given to provide a second chance for learners to correct their errors, i.e. eliminate incorrect response and figure out what the correct response should be.

Dependent Variables

The dependent variables were immediate posttest scores, delayed posttest scores and attitude questionnaire scores. Both immediate and posttest test scores involved : (a) overall correct responses on a posttest, (b) conditional probability of a correct response on an item on a posttest given that an incorrect response was made on a similar item during practice, and (c) conditional probability of an incorrect response on an item on a posttest given that an incorrect response was made on a similar item during practice. Attitude questionnaire scores involved a rating of learners' attitude toward the instruction, as measured by an attitude questionnaire.

Design and Treatments

The effects of two independent variables on three dependent variables were investigated through using a 2 * 3 factorial design. There were six treatments in this study, and among these treatments different types of feedback and a second try opportunity were given only after incorrect responses. Specifically, treatment 1, 2, and 3 provided different types of feedback (KOR, KCR or explanatory feedback), but all of them provided only one try for a missed practice question. Treatment 4, 5, 6 provided different type of feedback (KOR, KCR or explanatory feedback), but all of them provided a second try for a missed practice question. In treatment one through treatment six, whenever a correct response was produced, learners received KOR feedback stating "You are right." and were not asked to do the question again.

Materials

The instructional material used in this study was a CAI lesson on BASIC programming language, which was originally designed by Spock (1987) and systematically modified by the experimenter so that it matched the treatments for this study. The CAI lesson consisted of three submodules: (a) a brief introduction on what makes up a BASIC program; (b) instruction on variables and constants with four practice questions; (c) instruction on LET, PRINT, and INPUT statements of BASIC and two examples and practice questions included in the instruction for each statement. At the end of submodule (c), five practice questions were provided, which related to the use of LET, INPUT, and PRINT statements simultaneously in a BASIC program. Feedback which represented the treatments were provided after all the practice items involved in the three submodules of the CAI lesson.

All the practice questions were multiple-choice questions. They actually questioned students about common misconceptions held by computer novices regarding the LET, PRINT, and INPUT statements in BASIC (Mayer, 1976). Each question contained five distractors with one correct answer. The student was informed to choose the single correct answer from the five distractors. However, all distractors were designed to be equally plausible, so that answering correctly would not be simply a matter of eliminating highly implausible distractors. All the questions were designed to be difficult enough so as to make feedback yield effect. If the practice questions were too easy, learners might never make mistakes. If this was the case, learners might not need any feedback information to confirm if they have reached the goal of learning or to correct their mistakes. In such a situation, feedback would become useless.

Assessment Instruments

The assessment instruments included: (a) a pretest of 15 questions for grouping subjects into high, medium, and low prior knowledge groups; (b) an immediate posttest of 15 questions to measure learning achievement; (c) a delayed posttest of 15 questions for evaluating retention; (d) a questionnaire for asking subjects' attitude towards the instruction.

Except for the questionnaire, questions contained in the pretest, immediate posttest and delayed posttest were multiple-choice questions with five distractors and one correct answer. The tests were parallel, rather than identified test forms. The test items were similar across the pretest, immediate posttest and delayed posttest, each item reflecting a specific learning objective. In this study, the reason for using multiple-choice questions was that, multiple-choice questions were functionally appropriate for assessing rule-learning performance because the distractors involved in the questions could be constructed in such a way that learners were required to respond to more than one concrete instance of the rule. In addition, the distractors could be designed in a way that they served as cues for activating the recall of the previously learned concepts that make up the rule. Also, the great advantage of computers in administering multiple choice questions was that each distractor could be composed to represent a common misconception regarding application of a particular rule or concept, and the feedback could then be constructed to address that particular misconception.

Content Validity

To ensure the content validity of the pretest and two posttests used in this study, all the test items were designed to sample a representative portion of the subject content. The content validity of the pretest and two posttests used in this study were established by the following steps:

1. Prior to the pilot study, all test items were first reviewed by a professor, who teaches BASIC programming language at a southwestern university. All test items then were revised or rephrased according to his suggestions after the pilot study.

2. Two content specialists (one Ph. D., one Ph. D. candidate) in the field of Instructional Technology were asked to rate the relevance of the revised test items to the objectives identified in the task analysis. Both specialists were asked to use a 4-point scale ranging from *"not relevant (1)", "somewhat relevant (2)",* " *"relevant (3)",* to *"very relevant (4)"*. The mean score for each item resulted from their rating was used as an index for the degree of content validity. The mean scores varied between values of 3.5 and 4.0 which put items in the relevant to very relevant category. After a pilot study, those items with lower ratings were revised for problems in language, content, and directions.

Test Reliability

The internal reliability was estimated based on the average correlation among the test items. The K-R 20 formula (Richardson & Kuder, 1939) was employed to compute the internal reliability through analyzing each test itemcontained in the immediate posttest during the pilot study. After the K-R 20 was performed, a coefficient alpha of 0.6047 was obtained. Unfortunately, due to experimenter's oversight, these data, plus additional item analyses data, were not used to revise the test prior to administration of the full study.

Hypotheses

Main Effects. It was expected on both the immediate and delayed posttests that explanatory feedback group would score better than either KCR or KOR feedback group and KCR feedback group would score better than KOR feedback group (explanatory feedback group > KCR feedback group > KOR feedback group) with respect to (a) overall correct responses; (b) the conditional probability of a correct response on a test given that an incorrect response was made during practice and (c) higher score on the attitude questionnaire.

With respect to the conditional probability of an incorrect response on a test given that an incorrect response was made during practice, it was expected that on both the immediate and delayed posttests that the explanatory feedback group would have a lower probability of repeating errors than either KCR or KOR feedback group and KCR feedback group would have a lower probability of repeating errors than KOR feedback group (explanatory feedback group < KCR feedback group < KOR feedback group). It was expected on both the immediate and delayed posttests that groups receiving a second try (treatment 4, 5, 6) would score better than those who received a single try (treatment 1, 2, 3) with respect to (a) overall correct responses; (b) the conditional probability of a correct response on a test given that an incorrect response was made during practice and (c) higher score on the attitude questionnaire.

Interaction Effects. It was expected on both the immediate and delayed posttests that participants receiving explanatory feedback coupled with a second try (treatment 6) would score better than all other groups and participants receiving KCR feedback coupled with a second try would score better than those who received KOR feedback in combination with a second try with respect to (a) overall correct responses; (b) the conditional probability of a correct response on a test given that an incorrect response was made during practice and (c) higher score on the attitude questionnaire.

With respect to the conditional probability of an incorrect response on both the immediate and delayed posttests given that an incorrect response was made during practice, it was expected that explanatory feedback with a second try group would score lowest and KCR feedback with a second try group would have a lower probability of repeating errors than KOR feedback with a second try group.

Procedure

This study was conducted in a normal class period across three class periods, in three weeks. This study was conducted with the following steps:

1. One week prior to the experiment, a pretest was administered to those who decided to participate. Participants then were grouped into high, medium, and low prior knowledge groups based on their pretest performance. Participants in each of these three groups were randomly assigned to one of the six treatments.

2. The following week, the lesson was administered. Participants in each of the six treatments received the instruction through a CAI module and took the

immediate posttest and the paper-and-pencil attitude questionnaire. All participants were expected to finish in 1 hour.

3. A delayed posttest was administered one week after the experiment had been conducted.

Data Analyses

Overall analyses. A series of analyses of variance on each dependent variable was performed to determine which of these variables are statistically significant. If a significant ANOVA F were found, a series of Tukey's post hoc tests were performed to determine which treatments differed significantly from one another. The level of significance (*p*) was set at 0.10 in making decision whether to accept or reject the null hypotheses.

Analyses of conditional probabilities. The analyses of conditional probabilities were performed in accordance with methodology developed by Kulhavy (1977):

1. The conditional probability of a correct response on a test (on the immediate test: **Rit,** on the delayed posttest: **Rdt**) given that an incorrect response was made during practice **(W1)** is the proportion consisting of the total of number of the responses which were correct on the test and were incorrect during practice, $(Rit \cap Wl)$ and $(Rdt \cap Wl)$, divided by the total number of incorrect responses made during practice. This conditional probability can be expressed by:

$$P(Rit / Wl) = \text{number of } (Rit \cap Wl) / \text{number of } W1$$

$$P(Rdt / Wl) = \text{number of } (Rdt \cap Wl) / \text{number of } W1$$

2. The conditional probability of an incorrect response on the test (on the immediate posttest: **Wit** and on the delayed posttest: **Wdt**) given that an incorrect response was made during practice, $P(Wit/Wl)$ and $P(Wdt/W1)$, is the proportion consisting of the total of the number of responses which were incorrect on the test and were also incorrect during practice, $(Wit \cap Wl)$ and $(Wdt \cap Wl)$, divided by the total number of incorrect responses made during practice. This conditional probability can be expressed by:

$$P(Wit / Wl) = \text{number of } (Wit \cap Wl) / \text{number of } W1$$

$$P(Wdt / Wl) = \text{number of } (Wdt \cap Wl) / \text{number of } W1$$

In order to test for statistical difference between the six treatment groups, the proportional data was transformed by using the arc sine function (Dowdy & Wearden,1983). All statistical analyses were performed by using the SAS programs on an IBM 3081.

Results

Descriptive Data

Seven measures were used in this study. A summary table of the means and standard deviations for all participants on each dependent variable is presented in Table 1, and a summary of the the means and standard deviations of each of the six treatments on each dependent variable is presented in Table 2 and 3.

(Insert Table 1, 2 and 3 about here)

Hypotheses for Main Effects: Feedback Factor

Overall correct responses on the two posttests. The results of ANOVA (Table 4) revealed a significant difference in the three feedback groups with respect to the overall correct responses on the immediate posttest, F (2, 152) = 2.47, $p < 0.10$. Tukey's post hoc test revealed a significant difference between the explanatory feedback group and the KOR feedback group, indicating that explanatory feedback was significantly superior to KOR feedback in producing correct responses on the immediate posttest. Although the results revealed that the pattern of means of the explanatory feedback group and the KCR feedback group were in the predicted direction, the difference was not significant. Also, the mean score of the KCR feedback group was higher than the KOR feedback group, however, the mean difference between these two groups was not significant either (see Table 2 and Table 3). With respect to overall correct responses on the delayed posttest, the results failed to reveal a significant difference between the three feedback groups, indicating that students performed similarly on the delayed posttest no matter what type of feedback they received during practice.

(Insert Table 4 and 5 about here)

Attitude toward the instruction. With respect to participants' attitude toward the instruction, the results failed to reveal any significant difference between the three different feedback groups, indicating that students had similar attitudes towards the instruction in spite of the type of feedback message they received during practice.

The analyses of P(Rit/Wl) and P(Rdt/Wl). Regarding the means of the probability of correcting an incorrect response during practice on the immediate and delayed posttest, the results did not reveal a significant difference between the three feedback groups. Such results indicated that the effects of KOR, KCR, or explanatory feedback were similar in increasing the probability of changing an initial practice error to a correct response on the immediate and delayed posttest.

The analyses of P(Wit/Wl) and P(Wdt/Wl). Regarding the means of the probability of repeating an initial practice error on the immediate posttest, the result did not reveal a significant difference between the three feedback groups. Such result indicated that the effects of KOR, KCR or explanatory feedback were similar in reducing the probability that an practice error would occur again in the immediate posttest. However, the means of the probability of repeating an initial practice error on the delayed posttest did reveal a significant difference between the three feedback groups, F (2, 125) = 0.03, $p < 0.10$. Tukey's post hoc test revealed that the difference between the explanatory feedback group and the KCR feedback group was significant, indicating that explanatory feedback was significantly superior to KCR feedback in reducing the probability of repeating a practice error on the delayed posttest.

Although the results revealed that the explanatory feedback group produced a lower probability of repeating errors than the KOR feedback group and the KCR group produced a lower probability of repeating errors than the KOR group, significant differences were not revealed by the post hoc test (see Table 2, 3 and 4).

Hypotheses for Main Effects: Tries Factor

Overall correct responses on the two posttests. Regarding the overall correct responses measures, the results of ANOVA failed to reveal a significant difference between the two tries groups either for the immediate posttest or for the delayed posttest. Such results indicated that learners produced similar posttest performance no matter how many tries, one or two, followed their errors during practice.

Attitude toward the instruction. With respect to learners attitude toward the instruction, the results did not reveal any significant difference between the two different tries groups, indicating that students had similar attitudes toward the tries factor whether they received a single or a second try during practice.

The analysis of P(Rit/WI) and P(Rdt/WI). Likewise, regarding the means of the probability of a correct response on the immediate posttest or the delayed posttest given that an incorrect response was made during practice, the results did not reveal a significant difference between the two tries groups. Such results indicated that the effects of one or two tries were similar in increasing the probability of changing an initial error to a correct response on both the immediate and delayed posttests.

The analysis of P(Wit/WI) and P(Wdt/WI). Regarding the means of the probability of repeating an initial practice error on the immediate or the delayed posttest, the result did not reveal a significant difference between the two tries groups. Such results indicated that the effects of one or two tries were similar in reducing the probability that an practice error would occur again on both the immediate and delayed posttests.

Interaction Effects

In the analyses of interaction effects, no significant interactive effects were found between levels of feedback and number of tries on all dependent variables involved in this study, indicating that number of tries was not differentially effective across levels of feedback on any of the dependent variables.

Summary

The results indicated that there was a significant difference in the explanatory feedback group and the KOR feedback group regarding the production of correct responses on the immediate posttest. The explanatory feedback showed superiority over the KOR feedback in raising the total number of correct responses on the immediate posttest. The results also indicated a significant difference between the explanatory feedback group and the KCR feedback group while measuring the conditional probability of reducing a practice error on the delayed posttest. The explanatory feedback also seemed to be better than the KCR feedback in avoiding a previous practice error on the delayed posttest. Other measures on the effects of the feedback factor failed to indicate any significant results, and no significant results were found from the measures for the tries factor or for the interaction effect.

Discussion and Recommendations

Main Effects: Feedback Factor

Explanation and discussion of the results. Among the findings discussed above, the first significant result: the superiority of explanatory feedback over KOR feedback on overall correct responses on the immediate posttest, seems to indicate that explanatory feedback supplies learners with the corrective information so that they can determine why a correct answer is right and why an

As error occurs, thereby producing more correct responses on the immediate posttest. for the second significant result, the superiority of explanatory feedback over KCR feedback on reducing the probability of repeating an error on the delayed posttest, it also seems to indicate that explanatory feedback allows learners to incorporate the corrective information into their ongoing information-processing procedure so that they will not repeat an previous error. This finding was particularly interesting in that it was found on the delayed posttest.

Although patterns of means were in the predicted direction on all measures examining effects of levels of feedback (explanatory feedback > KCR feedback > KOR feedback), the significance of these differences were inconsistently significant. Possible explanations for these effects are twofold: (a) the learning motivation was suppressed by the difficulty of the learning materials and (b) low reliability of the testing instrument interjected more error variance. As matter of fact, the grand mean across all treatment groups was 7.24 for the immediate posttest, 6.25 for the delayed posttest, with total possible scores of 15 for each measure. So the average performance across all treatment groups was quite low. This indicated the level of the difficulty and complexity of the instruction, practice questions and test items may have been beyond the ability level of most participants. Students may not have maintained the level of attention and motivation expected from them throughout the whole instruction. It is necessary to point out that the instructional materials used in the current study were intended to be written at a relatively difficult level in order to ensure some errors would occur during practice so that feedback could be utilized.

It is also necessary to emphasize that according to Polya (1957), successful problem solving behavior requires stages of understanding the problem, devising a plan, carrying out the plan and evaluating the effectiveness of the plan. He further cites the importance of the effective aspects of problem solving such as motivation, determination and curiosity in problem solving performance. In this study, if learners had difficulty in understanding or acquiring the learning information, they might not have been fully motivated to complete the whole lesson or to exercise any of the four problem solving stages in order to successfully answer the practice questions. If so, explanatory feedback given after a practice error may have worked as a remedial instruction. In such a case, if a test was administered immediately after the lesson, those who received explanatory feedback during practice may still have been able to recall the corrective information to produce the correct response on that test. In the same case, those who received KCR feedback may simply have memorized the correct answer and matched that answer with an answer similar in form. As for learners in the KOR feedback group, if they were unmotivated or unable to learn, a simple message such as "You are right" or "You are wrong" may have been insufficient to aid them to analyze an error or to reason out the correct answer during practice, requiring more mental effort than they were able or willing to apply. This may have resulted in poor performance on the immediate posttest. These reasons may be responsible for the finding that a significant difference was only found in overall correct responses on the immediate posttest between the explanatory and the KOR feedback groups. And these reasons may also be responsible for why the mean difference of the overall correct responses on the immediate posttest between the explanatory and the KCR feedback groups was in the predicted direction but not great enough to reach the statistical difference.

As for why explanatory feedback was significantly better than KCR feedback in reducing the probability of repeating a practice error on the delayed posttest, the reasons may be that if students did not maintain a high level of attention or motivation throughout the entire CAI lesson, they might have been unable to employ their learning to answer the practice questions and might have missed most of the questions. If this were the case, the more practice errors, the more explanatory information would be received by students in the explanatory feedback group. In the same case, a number of correct answers would be received by students in the KCR feedback group and the message containing fairly strong language such as "You are wrong" would be repeatedly received by students in the KOR feedback group. As a result, when the delayed posttest was administered, students received explanatory feedback might not have been able to apply their learning to determine the correct answer, however, they might have been able to employ the explanatory feedback information sufficiently to avoid the previous error. Those who in the KOR feedback group might not commit the same previous error because they were told "You are wrong" before.
As for learners in the KCR feedback group, because they were only provided with the correct answer, they might have easily forgotten the correct answers due to their inability to understand why the correct answer was right. This reason might have been responsible for why, with respect to the measure, probability of repeating a practice error on the delayed posttest, the explanatory feedback group scored significantly lowered than the KCR feedback group and the KOR feedback group scored slightly lower than the KCR feedback group (the difference between the KCR and KOR groups was not great enough to be significant). Again, if students were not motivated to learn or had difficulty in comprehending the materials, whatever they read from the CAI lesson would have been faded quickly from their memories. As a result, all the three feedback groups might have performed poorly and similarly on a posttest that was delayed. This reason might have been responsible for the failure to reach a significant difference on the measures of overall correct responses and the probability of changing a practice error to a correct response on the delayed posttest among the three feedback groups. It might be that because students more often had failure rather than success with the learning materials, they might have had less positive attitudes toward the instruction in spite of the different treatments they received. On the measure of students' attitude toward the instruction, the mean difference among the three feedback groups was not statistically significant, however, in terms of mean differences, students in the explanatory feedback group scored higher on the attitude questionnaire than did students in the KCR and KOR feedback groups, and students in the KCR feedback group scored higher on the attitude questionnaire than did those who in the KOR feedback group.

The above assumptions can be supported by the following evidences. The subjects involved in this study were students enrolled in the Computer Literacy course. Most students take the Computer Literacy course because it is required by their departments. Consequently, many important characteristics such as need for achievement, level of interest, attention and motivation might be different in other populations. Besides, results obtained from the pretest revealed that ninety percent of participants were considered as low prior knowledge students. From analyzing students' preassessment information, it was found that most of the participants had never have any experience with computers. It may be that because most participants did not have any prior knowledge on computers and programming languages, their learning motivation might have been seriously suppressed. They might have even been anxious about learning this content, further depressing their motivation and performance. Comments made on the attitude questionnaire confirmed the above assumption. These comments indicated that students felt that both the instructional materials and the practice questions were fairly hard. Those comments also revealed that students had difficulty in learning from the instructional materials and seemed to have troubles in completing the practice questions during the class time. The test items might have been too hard for most participants, and there might be some inadequacy in the test item for the reliability analysis of this instrument revealed a fairly low reliability coeffcient (0.6047).

Recommendations for future studies. Based upon the above discussion, a number of recommendations are made: first of all, on one hand, differences which existed among the sample such as level of learning motivation, ability and need for achievement should be seriously taken into consideration when investigating the effects of different types of feedback. On the other hand, future studies are recommended to examine how the level of difficulty of instructional materials affects the learning motivation and, consequently the facilitating effects of different types of feedback. Think-aloud interviews are recommended so that the researcher can obtain information as to how learners utilize the feedback information. Self-report measures of effort and motivation might also be used to obtain information to substantiate the above assumptions .

Secondly, future research is necessary to determine what is the sufficient and adequate amount of feedback information that would be beneficial for learning. Practically, it is recommended that instructional designers reconstruct explanatory feedback into two specific types: (a) the corrective feedback, which contains corrective information and is primarily for helping the learner to locate an error and (b) the remedial feedback, which contains brief instruction as to hint the learner to search for the correct answer. Also recommended is that both the corrective and remedial feedback should be included in any CAI lesson, and learners should be allowed to choose the specific type of feedback they really need. Of course, researchers should examine the comparative benefits of the corrective and remedial feedback on the effects of (a) correcting an initial error to a correct response on a later performance, (b) maintaining an initial correct response on a later performance, and (c) decreasing the probability of repeating an initial error on a later performance.

Main Effect: Tries Factor

Explanation and discussion of the results. Unexpectedly, with respect to all dependent variables involved in this study, the results failed to support the hypotheses that a second try upon a missed practice practice would be more effective than a single try. As a matter of fact, only means of the attitude toward the instruction were in the predicted direction.

One of the possible reasons for these unexpected results might be that, as discussed, students involved in this study might have been unmotivated to maintain the level of attention expected from them throughout the CAI learning. Students might have passively gone through the whole CAI lesson without paying attention to the learning content. When they were asked to do the practice questions, they might have experienced frustration due to their inability to understand the questions. They might have expected to finish the practice questions as soon as possible. If so, during the first try, learners might have no idea about how to find the correct answer. After an error, feedback might have been insufficiently remedial to correct misconceptions of rules and concepts. Thus, during the second try, even with the help of feedback, they might not have been able to eliminate the previous error or to determine the correct answer.

Very possibly, students might have been seriously confused by the feedback information and then might have made an extra error during the second try. In this case, students receiving a second try after an error might have resulted in worse posttest performance than those who received a single try. Those who received a single try after after an error did not have an additional opportunity to make an extra error, therefore this might have reduced the probability of repeating a previous error on a later performance. These findings were consistent with the findings obtained in Hodes'(1984) study in which an extra corrective opportunity upon the missed item was not superior to a single try upon the same item in raising the posttest scores.

Recommendations for future studies. Future studies are needed to determine the level of effort and motivation that might be necessary for second try to be effective. The question of whether the instructional event such as a second try provided during learning or practicing is necessary should be answered. Further studies investigating the effects of second try and its pairing with feedback should be conducted to ascertain if this study's results are generalizable to other populations and other content. These future studies should include some important conditional probability measures, which were not included in this study. The suggested conditional probability measures may include the probabilities of (a) maintaining a correct response made during a second try on a later performance and (b) repeating an error made during a second try on a later performance. These measures can provide information as to whether a second try can help learners to change a previous error to a correct response and to maintain that correct response on a later performance.

Interaction Effects

Explanation and discussion of the results. In the analyses of interaction effects, results failed to reveal any significant results. One of the possible reasons for this might be a result of the unexpected findings obtained from the analysis of the tries factor. In the analysis of the tries factor, only the means of the attitude toward the instruction were in the predicted directions. Without the predicted effects of the tries factor, the interaction would not be expected.

Recommendations for future studies. The major expectation of this study was that a combination of explanatory feedback with a second try would provide learners with corrective information as well as an extra opportunity to review the feedback information and to redo the missed question, thereby producing a superior posttest performance. The findings did not support this prediction. Future studies are necessary to determine how much and what kind of corrective information paired with a second try can really make a difference compared to simple KOR or KCR feedback paired with a second try on error correction. Future research is also needed to investigate when and why feedback paired with a second try should be given when considering other important variables such as learning motivation and other learners' characteristics.

Conclusion

In the analysis of feedback factor, the findings indicated that explanatory feedback was more powerful than KOR feedback in helping students maintain the correct responses on the immediate posttest. Explanatory feedback was also more powerful than KCR feedback in preventing students from repeating a previous error on a delayed posttest. No significant results were obtained either from the analyses of the tries factor or for the interaction effects.

One of the possible reasons for these spotty and inconsistent results might be that learners did not maintain a high level of learning motivation. This low attention and motivation level might have been due to the level of difficulty and complexity of the instruction which seemed to be beyond the ability level of most participants. The findings from this study indicate future research is necessary on determining how learners utilize feedback information and how much information is sufficient and appropriate to encourage a change in errors into correct responses. Future study is also needed to determine how the difficulty of instruction, the level of effort and motivation and learners' characteristics influence the effectiveness of different types of feedback and their pairing with a second try on error correction.

References

Ausubel, D. P. (1968). Educational psychology: A cognitive view, New York:Holt, Rinehart and Winston, Inc.

Buss, A. H., & Buss, E. (1956). The effect of verbal reinforcement combinations on conceptual learning, Journal of Experimental Psychology, 52, 283-287.

Carter, J. (1984). Instructional learner feedback: A literature review with implications for software development, The Computing Teachers , 12(2), 53-55.

Dowdy, S. & Wearden, S. (1983). Statistics for research, p. 299-301, New York:John Wiley & Sons, Inc.

Gagne, R. M., & Briggs, L. J. & Wager (1987). Principles of instructional design (2nd ed.) New York:Holt, Rinehart and Winston.

Gilman, D. A. (1969). Comparison of several feedback methods for correcting errors by computer assisted instruction, Journal of Educational Psychology, 60, 503-508.

Hodes, C. L. (1984-1985). Relative effectiveness of corrective and noncorrective feedback in computer assisted instruction on learning and achievement. Journal of Educational Technology System, 1(4), 249-254.

Hudson, R. C. (1981). Feedback: Administered by a computer, Dissertation Abstracts International, 42, abstract no. 4420A.

Kulhavy, R. W. (1977). Feedback in written instruction, Review of Educational Research,47, (1), 211-232.

Mayer, R. E. (1976). Some conditions of meaningful learning for computer programming: advanced organizer and subject control of frame order. Journal of Educational Psychology, 68, No. 2, 143-150.

Moore, J. W., & Smith, W. I. (1965). Role of knowledge of results in programmed instruction. Psychological Reports, 14, 407-423.

Park, 0. (1984). Empirically Based Procedures for Designing a Response-Sensitive Sequence in Computer-Based Instruction: An Example From Concept-Teaching Strategies. Journal of Computer-Based Instruction, 11(1), 14-18.

Phye, G. D. (1979). The Processing of Informative Feedback about Multiple-Choice Test Performance, Contemporary Educational Psychology, 4, 381-394.

Polya, G. (1957). How to Solve it (2nd ed.), New York: Doableday.

Roberts, F. C., Park, 0. (1984). Feedback strategies and cognitive style in computer-based instruction, The Journal of Instructional Psychology, vol. 11, No.2.

Roper, W. J. (1977). Feedback in computer assisted instruction. Programmed Learningand Educational Technology, 14, 43-49.

Ross, R. W. (1984). Matching the lesson to the student: Alternative adaptive designs for individualized learning systems, Journal of Computer-Based Instruction, 11, 42-48.

Ross, 5. M., Rakow, E. A. et al (1980). Instructional adaption for selfmanaged learning systems, Journal of Computer-based instruction, 11, 42-48.

Saloman, G. (1972). Heuristic models for the generation of aptitude treatment interaction hypotheses. Review of Educational Research, 42 (3), 327-343.

Sassenrath, J. M. (1975). Theory and results on feedback and retention. Journal of Educational Psychology, 67, 894-899.

Skinner, B. F. (1961). Why we need teaching machines. Harvard Educational Review, 31 377-398.

Skinner, B. F. (1968). The Technology of Teaching, New York:Appleton Century-Crofts making.Educational Technology, 19(1), 20-26.

Spock, P. (1988). The Effect of feedback, correctness of response, and confidence of response on a rule-learning task using computer-assisted instruction, Doctoral Dissertation, the University of Texas at Austin, Austin, TX.

Sternberg, R. J. (1983). Criteria for intellectual skills training. Educational Researcher, 12, 6-12.

Table 1: Summary of Means and Standard Deviations for All Dependent Variables

Dependent Variables	N	M	SD
Immediate Posttest	158	7.244[a]	2.83
Delayed Posttest	132	6.25[b]	3.30
Attitude Questionnaire	160	30.06[c]	5.31
P(Rit/W1)	158	0.79	0.32
P(Wit/W1)	158	0.77	0.32
P(Rdt/W1)	110	0.57	0.36
P(Wdt/W1)	131	1.00	0.35

[a] Immediate posttest contained 15 multiple choice questions.

[b] Delayed posttest contained 15 multiple choice questions.

[c] Attitude questionnaire contained 10 questions (highest possible score was 50; higher score indicates more favorable attitude).

Table 2: Summary of Means and Standard Deviations of Six Treatments for Immediate Posttest, Delayed Posttest and Attitude Questionnaire

		Level	of	Feedback
Number of Tries		KOR	KCR	Explanatory
One	Immediate Posttest	N = 27 M = 6.40 SD = 2.35	N = 27 M = 7.74 SD = 2.39	N = 25 M = 7.92 SD = 3.23
	Delayed Posttest	N = 24 M = 5.91 SD = 2.74	N = 20 M = 7.15 SD = 3.37	N = 20 M = 6.90 SD = 4.15
	Attitude Questionnaire	N = 28 M = 29.28 SD = 4.67	N = 27 M = 29.55 SD = 6.29	N = 25 M = 30.72 SD = 4.65
Two	Immediate Posttest	N = 25 M = 6.72 SD = 2.63	N = 29 M = 6.62 SD = 2.96	N = 25 M = 7.64 SD = 2.88
	Delayed Posttest	N = 22 M = 5.77 SD = 3.20	N = 24 M = 5.08 SD = 2.63	N = 22 M = 6.95 SD = 3.44
	Attitude Questionnaire	N = 25 M = 28.72 SD = 4.56	N = 29 M = 29.96 SD = 5.90	N = 26 M = 31.26 SD = 5.82

Table 3: Summary of Means and Standard Deviations of Six Treatments for All Conditional Probability Measures

Level of Feedback

Number of Tries		KOR	KCR	Explanatory
One	P(Rit/W1)	N = 27 M = 0.75 SD = 0.26	N = 27 M = 0.86 SD = 0.33	N = 25 M = 0.81 SD = 0.36
	P(Rdt/W1)	N = 27 M = 0.54 SD = 0.23	N = 27 M = 0.57 SD = 0.39	N = 25 M = 0.60 SD = 0.47
	P(Wit/W1)	N = 24 M = 0.79 SD = 0.27	N = 20 M = 0.70 SD = 0.33	N = 20 M = 0.75 SD = 0.36
	P(Wdt/W1)	N = 24 M = 1.03 SD = 0.23	N = 20 M = 0.99 SD = 0.39	N = 20 M = 0.96 SD = 0.47
Two	P(Rit/W1)	N = 25 M = 0.78 SD = 0.30	N = 29 M = 0.73 SD = 0.24	N = 25 M = 0.81 SD = 0.36
	P(Rdt/W1)	N = 25 M = 0.53 SD = 0.33	N = 29 M = 0.47 SD = 0.34	N = 25 M = 0.60 SD = 0.47
	P(Wit/W1)	N = 20 M = 0.78 SD = 0.30	N = 24 M = 0.83 SD = 0.23	N = 21 M = 0.75 SD = 0.36
	P(Wdt/W1)	N = 20 M = 0.96 SD = 0.33	N = 24 M = 1.14 SD = 0.24	N = 21 M = 0.96 SD = 0.47

Table 4: Summary of Analyses of Variance for All Dependent Variables

Variables	F	df	p
Types of Feedback (F)			
Immediate Posttest	2.47	2, 152	0.080*
Delayed Posttest	1.29	2, 126	0.279
Attitude Questionnaire	1.81	2, 154	0.167
P(Rit/W1)	0.01	2, 152	0.909
P(Wit/W1)	0.09	2, 152	0.917
P(Rdt/W1)	1.98	2, 125	0.142
P(Wdt/W1)	2.58	2, 125	0.079*
Numbers of Tries (T)			
Immediate Posttest	0.68	1, 152	0.411
Delayed Posttest	−19	1, 126	0.209
Attitude Questionnaire	0.02	1, 154	0.877
P(Rit/W1)	0.106	1, 152	0.305
P(Wit/W1)	1.19	1, 152	0.277
P(Rdt/W1)	0.00	1, 125	0.949
P(Wdt/W1)	0.03	1, 125	0.852
F * T			
Immediate Posttest	0.95	2, 152	0.401
Delayed Posttest	1.40	2, 126	0.250
Attitude Questionnaire	0.17	2, 154	0.845
P(Rit/W1)	0.95	2, 152	0.388
P(Wit/W1)	0.65	2, 152	0.525
P(Rdt/W1)	0.94	2, 125	0.395
P(Wdt/W1)	1.64	2, 152	0.198

* $p < 0.10$

Table 5: Tukey's Post hoc test for Feedback Groups on the Immediate Posttest

	Differences	between	Means
Variables	KOR Feedback	KCR Feedback	Explanatory Feedback
KOR Feedback	0	0.60	1.22*
KCR Feedback	0.60	0	0.61
Explanatory Feedback	1.22*	0.61	0

* $p < 0.10$

Title:

Metacognitive and Cognitive Effects of Different Loci of Instructional Control

Author:

Miheon Jo Lee

METACOGNITIVE AND COGNITIVE EFFECTS OF DIFFERENT LOCI OF INSTRUCTIONAL CONTROL

With the adoption of microcomputers for education to be used as teaching or learning tools, the design of Computer-Assisted Instructional (CAI) programs has received more emphasis, particularly as such adoption of microcomputers allows for the interaction between learners and instructional materials. One of the issues currently attracting researchers' attention in the area of design of interactive CAI programs is the locus of instructional control: should control of instruction be given to a program (program control) or to a learner (learner control)?

Program control has been the dominant instructional design strategy, utilizing prescribed rules with little or no control given to learners over instructional components. The learner control approach challenges the dominant design strategy by avoiding a rigidly prescribed structure for instruction and emphasizing learners' own decision making in progressing through instruction.

Advocates of the learner control design strategy have discussed its various advantages. Of these possible benefits of learner control in comparison with program control, two were the focus in this study: 1) positive influence on metacognitive and cognitive knowledge and skills, and 2) positive approach to individual differences among learners.

First, one of the most important benefits of learner control is its positive effects on the enhancement of metacognitive and cognitive knowledge and skills through dynamic learning processes. Unlike the program control design strategy, which aims to improve only products of learning, the learner control strategy emphasizes processes as well as products of learning. Under learner control, individual learners are in a position to analyze the nature of a task; to monitor their own understanding of the task; to select activities to be performed; and to decide levels or amounts of instruction needed to extend the present state of knowledge, while continuously monitoring their own cognitive processes and performances. By being aware of control options and deliberately making decisions about the use of the options under learner control, learners can continuously evaluate their progress towards a given objective and attempt to develop effective learning strategies to achieve that objective. As a result, through learner controlled instruction learners may be able to integrate new learning materials more thoroughly within their existing knowledge schema, and enhance their thinking skills in a more effective way than they may through program controlled instruction.

A second potential advantage of learner control is its positive approach to individual differences among learners. Learner control as an instructional design strategy provides a learner with freedom to follow her/his own interests and needs, and to make decisions about learning rather than being locked into instruction predesigned by an instructor or an instructional designer. In comparison with program controlled instruction, in which a program is predesigned to have control in providing an instructional path best for learners, learner controlled instruction provides more flexible instructional path about which learners can make decisions. Flexibility is the intended outcome of the learner control design strategy. While a dynamic on-going adaptation to an individual learner's varying needs and interests is almost impossible to preplan for program controlled instruction, flexibility provided by learner controlled instruction may make such on-going adaptation possible, and thus interact positively with individual differences.

However, these intuitively appealing theoretical assumptions regarding learner control have received little support from actual research. Although significant efforts have been devoted to research on the locus of instructional control, the results have been diverse, making generalization difficult (Carrier, 1984; Steinberg, 1977). In addition, no previous research on learner control has addressed its metacognitive effects. Therefore, the theoretical assumptions remain as empirical issues requiring research attention.

Purpose

The purpose of the study was to investigate different loci of instructional control and prior background knowledge in terms of two aspects of metacognitive effects (utilization and correctness of metacognitive monitoring) and two of cognitive effects (knowledge acquisition and application). Prior background knowledge refers to the prior knowledge of geometric concepts (distance, direction and pattern) related to the LOGO commands taught in the study.

Three major research goals were established. The first was to investigate the loci of instructional control in terms of the two aspects of metacognitive effects and the two of cognitive effects. In assessing the effects on utilization and correctness of metacognitive monitoring, the Metacomponential Interview developed by Clements and Nastasi (1988) was adopted, focusing on one of the metacomponents, metacognitive monitoring of solution processes. For the assessment of the effects on knowledge acquisition and application, two versions of computer-delivered tests were developed by the researcher.

The second goal was to extend the investigation in relation to the effectiveness of learner strategies used under learner control. The effectiveness of learner strategies was evaluated on the basis of the following two criteria as recommended by Merrill (1984): accuracy of self-estimates of understanding and appropriateness of control use (Appendix D). To assess the accuracy of self-estimates, at the end of the tutorial session of each lesson students were informed that four practice questions would be presented, and asked to estimate how many of them they could get right. Then, the estimate was compared with the actual number of correct answers in subsequent practice. Concerning the assessment of appropriateness of control use, the frequency of a student's actual uses of control options was recorded by the learner control program, and it was compared with the student's performance in subsequent practice. On the basis of the median of the scores for effectiveness of learner strategy, the total learner control group was partitioned into two subgroups (effective and ineffective learner control subgroups). Comparisons were then made among the program control group and the two learner control subgroups concerning the effects on utilization and correctness of metacognitive monitoring, and on knowledge acquisition and application.

The third goal was to extend the investigation in relation to students' prior background knowledge, as one aspect of individual differences among students. The outcomes of each of the assessments on the two aspects of metacognitive effects and the two of cognitive effects were regressed for each of the loci of instructional control groups on scores of the pretest assessing prior background knowledge. Subsequently, the way in which the outcomes were regressed were compared between the learner control and program control groups.

Research Methods

Learning Materials

LOGO, a computer programming language, was the content area taught. For each of the loci of instructional control groups (program control and learner control groups), two parallel versions of CAI programs were developed and used for the study. The programs were delivered by Apple IIe computers.

Students in both program control and learner control groups were introduced to five basic LOGO commands (i.e., FORWARD, BACK, RIGHT, LEFT and REPEAT) with definitions, rules, examples, and practice. The two versions of the programs were the same in the tutorial session of each lesson, but differed in the presentation of the question for self-estimates of understanding and in the choices available over the use of relevant instructional components in the practice session.

The program for the learner control group consisted of an internally controlled CAI program in which students were free to decide whether and how often they needed to review some instructional components relevant to a practice item with the option of skipping the review. In contrast, the program control group was given an externally controlled CAI program in which all control in the lesson was regulated by the program on the basis of predesigned rules except the instructional pace -- students could advance through instruction when they were ready by pressing the space bar or the return key.

In attempting to find ways to overcome some problems found in other research on the locus of instructional control, the following factors were integrated into the design of instruction and research, as recommended by researchers and theorists. Such integration of supportive factors seemed to be especially important for the learner control group whose time to be devoted to the task was more in their hands than for the program control group. First, in order to motivate students and simulate an actual classroom-like incentive, students were informed that although scores on practice and tests would not impact their class grades, their teachers would look at their scores (Tennyson, 1980; Tennyson and Buttrey, 1980). Second, students were required to solve a certain number of practice items to assure that they did not bypass important instructional components (Kinzie, 1988; Steinberg, 1977; Waldrop, 1984). Third, prior to the instruction, there was a pretraining period to help students become familiar with computers and the novel learning system, perform conscious cognitive information processing, and understand the objectives, procedures and values involved in building their own learning strategies (Gay, 1986; Hannafin, 1984; Kinzie, 1988; Merrill, 1984; Schmitt and Newby, 1986; Steinberg, 1977).

In addition, the following two factors were integrated into the design of learner controlled instruction as recommended by other researchers and theorists. Control options were clearly labeled and presented in a menu format for easy use of the options (Reigeluth and Stein, 1983). Students were asked to estimate how many of the subsequent practice questions they would get right so that they could be specifically guided to evaluate their own understanding of the given learning material (Garhart and Hannafin, 1986; Merrill, 1984). Also, immediate feedback, continuous advice on students' on-going progress and summaries of their uses of control options were presented in order to help students make informed decisions about their own learning (Hannafin, 1984; Kinzie, 1988; Lepper and Malone, 1987; Markman, 1977; Merrill, 1984; Tennyson, 1980; Tennyson and Buttrey, 1980; Waldrop, 1984).

Subjects

The subjects of the study were 62 third graders enrolled in a suburban elementary school in Wisconsin during the spring of 1990. None of the students participating in the study had prior experience with LOGO.

Procedures

The study was conducted in the school microcomputer lab over seven weeks. When students arrived at the lab, they were seated in front of the computers that showed their names on the screen and contained the instructional program versions assigned to them. Students worked with computers individually. Each participating student took lessons and assessments through eight sessions, and had an interview.

In the first session, a pretest was administered to assess students' prior background knowledge (Appendix B). Then, before instruction began, on the basis of the pretest scores students were randomly assigned to either the learner control or the program control group. To keep the levels of students' prior background knowledge as equal as possible, a Yoked Randomized Assignment method (Cohen, 1969) was utilized. As the result, 31 students were assigned to each locus of instructional control group.

The next five sessions were used for lessons on LOGO commands. Before the first lesson began, a short introductory training period was given to students to help them become familiar with computers and instructional programs and understand the objectives, procedures and values related to the use of the programs. In the sessions for lessons, students who finished a lesson earlier than the others were asked to remain in the lab and spend the rest of the session with a paper and pencil maze task in order to encourage students to devote adequate efforts to their given computer tasks.

Upon the completion of all the lessons, the seventh and eighth sessions were used for posttests on knowledge acquisition and application. For the assessment of knowledge acquisition, a 20-item posttest was administered to students. In the test, items similar to the practice ones embedded in lessons were presented to evaluate students' abilities to recall the LOGO commands taught throughout the study. The number of test items on each LOGO command was equally balanced across all the commands taught. For the assessment of knowledge application, given the length of each side of a square on the screen, students were asked to draw a square by entering commands for turtle moves and turns. Both tests were delivered by computers.

The interview on metacognitive monitoring was conducted with individual students after all the lessons were over. Four questions were asked, with five prompts for each question (Appendix C). The entire interview process was recorded on audio tape for later analysis and testing of inter-coder agreement on scoring. When the inter-coder agreement was tested, the Cronbach's α reliability values were .8700 for the utilization of metacognitive monitoring, and .9864 for the correctness of metacognitive monitoring.

Hypotheses

The following hypotheses were examined in relation to the three respective research goals. For the first and third goals, the hypotheses were made for confirmatory purposes. In contrast, for the second goal, the hypothesis was made for an exploratory purpose.

1) There will be differences in means between the program control group and the learner control group, in favor of the latter group, in the assessments of two aspects of metacognitive effects (utilization and correctness of metacognitive monitoring) and two of cognitive effects (knowledge acquisition and application).

2) There will be differences in means among the program control group and the two learner control subgroups partitioned according to the effectiveness of learner strategy, in the assessments of utilization and correctness of metacognitive monitoring and knowledge acquisition and application.

3) There will be differences between the program control group and the learner control group in the ways in which the outcomes of each assessment are regressed on the scores of a test assessing prior background knowledge; the slopes of the regressed outcomes of each assessment will be steeper for the program control group than for the learner control group.

Data Analyses and Results

Data analyses and their results are discussed below, in response to the hypotheses made for each of the three research goals.

Investigation of Metacognitive and Cognitive Effects of the Loci of Instructional Control

In serving the first goal of the study, the program control group was compared with the learner control group in the assessments of the two aspects of metacognitive effects and the two of cognitive effects.

A one-tailed t test was employed in comparing mean scores of the two groups in those assessments (Tables 1 and 2). Significant differences between the two groups were found in favor of the learner control group in all the assessments. Thus, the results of the t test supported the first research hypothesis.

Insert Table 1 and Table 2 about here

Extended Investigation in Relation to Effectiveness of Learner Strategy

For the second goal of the study, the investigation of the loci of instructional control was extended in relation to the effectiveness of learner strategy used under learner control. A descriptive data analysis was conducted on the accuracy of self-estimates of understanding, appropriateness of control use and effectiveness of learner strategy (Tables 3, 4 and 5). The analysis showed similar distributions for accuracy, appropriateness and effectiveness across possible ranges of scores.

Insert Table 3, Table 4 and Table 5 about here

To permit a further investigation of the metacognitive and cognitive effects of the two loci of instructional control, the learner control group was partitioned into two subgroups (effective and ineffective learner control subgroups) according to the effectiveness of learner strategy. Then a one-way analysis of variance (ANOVA) was conducted to compare mean scores of the two learner control subgroups and the program control group in the assessments of the two aspects of metacognitive effects and the two of cognitive effects (Tables 6 and 7). As hypothesized, concerning the metacognitive effects, significant differences were found among the three groups in the assessments of both aspects: utilization and correctness of metacognitive monitoring. However, concerning the cognitive effects, a significant difference was found in the assessment of one aspect, knowledge acquisition. In the assessment of knowledge application, although a statistically significant difference was not found, the learner control subgroups, especially the effective learner control subgroup, demonstrated better performance than the program control group.

Insert Table 6 and Table 7 about here

In addition to the ANOVA, to determine where significant differences were, Fisher's least significant difference test was employed for multiple pairwise comparisons among the three groups at the familywise $\alpha=.05$ level (Table 8). The comparisons were made with the mean scores of each of the assessments for the three dependent variables for which significant differences were found from the ANOVA: utilization of metacognitive monitoring, correctness of metacognitive monitoring, and knowledge acquisition. The comparisons between the effective learner control subgroup and the program control group revealed that the former group demonstrated significantly better performance than the latter group in the assessments of all of the three dependent variables. Between the ineffective learner control subgroup and the program control group, significant differences were found in favor of the former group only in the assessments of the two variables measuring metacognitive effects, namely, utilization of metacognitive monitoring and correctness of metacognitive monitoring. Furthermore, between the two learner control subgroups, no significant difference was found in any assessment.

Insert Table 8 about here

Extended Investigation in Relation to Students' Prior Background Knowledge

In serving the third goal of the study, the investigation of the metacognitive and cognitive effects of the loci of instructional control was extended in relation to learners' prior background knowledge.

A regression analysis was used to test the relative contribution of students' prior background knowledge to each of the assessments of the two aspects of metacognitive and the two of cognitive effects. For each locus of instructional control group, the outcome of each of the assessments was regressed on the scores of a pretest assessing prior background knowledge (Table 9).

Insert Table 9 about here

Then, to test whether the regression lines of the two loci of instructional control groups were parallel for each assessment, equality of the regression slopes for the two groups was tested according to standardized procedures (Finn, 1974, 376-378; Rao, 1973, 281-284). The regression slopes turned out to be steeper for the program control group than for the learner control group for each of the assessments. However, the results of the test indicated that the slopes of the regression lines on each assessment were statistically the same for the learner control and the program control groups: F(1,58) = 0.0658, n.s. for utilization of metacognitive monitoring; F(1,58) = 1.4702, n.s. for correctness of metacognitive monitoring; F(1,58) = 0.4181, n.s. for knowledge acquisition; and F(1,58) = 1.4083, n.s. for knowledge application. Thus, the third research hypothesis formulated in this study was not supported by these results.

Discussion

In attempting to find ways to overcome problems that students, especially those with relatively low levels of prior background knowledge, may have under learner control, supportive factors such as provision of clearly labeled options, basic requirements, and presentation of feedback and advice on on-going progress were integrated into the design of instruction. When these supportive design elements were integrated into the design of a learner control study, learner control seemed to foster students' metacognitive as well as cognitive knowledge and skills in a more effective way than did program control. The overall performance of the students under learner control was significantly better than that under program control in the assessments of two aspects of metacognitive effects (utilization and correctness of metacognitive monitoring) and two of cognitive effects (knowledge acquisition and application).

In fact, under learner control, learners become responsible for their own learning, practice conscious reflection on their cognitive abilities, task demands and learning strategies, and learn how to manage their own thinking processes and learning activities. Thus, in those processes, a learner can be led to enhance her/his metacognitive monitoring skills to perform and behave strategically. At the same time, in such learning processes more cognitive burden is given to a learner and thus s/he can improve her/his ability to acquire knowledge more thoroughly and to apply the acquired knowledge to other tasks.

It has been stressed that in addition to the concern with teaching content-specific skills and knowledge, educators need to pay attention to teaching of content-general skills and knowledge requiring high level thinking (Bonner, 1988; Chipman and Segal, 1985; Gagne and Glaser, 1987; Perkins and Salomon, 1989; Schmitt and Newby, 1986). Metacognitive skills and knowledge are one area that needs such attention. On this point, metacognitive skills and knowledge do not automatically improve as age increases, so an attempt must be made to provide learners with opportunities to learn and practice such skills and knowledge (Brown, 1980; Markman, 1977; Schmitt and Newby, 1986). The present study shows the potential of learner control as an instructional design strategy that may provide learners with opportunities to improve metacognitive skills and knowledge, while attempting to teach skills and knowledge of specific subject matter content.

However, the pairwise comparisons between each of the learner control subgroups and the program control group indicate the crucial need for accurate self-estimates of understanding and appropriate uses of control options to achieve the full potential of learner control in enhancing both metacognitive and cognitive skills and knowledge. In spite of the integration of the supportive factors into the design of instruction, there was still a wide range in the distribution of data on accuracy of self-estimates of understanding, appropriateness of control use and effectiveness of learner strategy.

Furthermore, concerning students' prior background knowledge, a post hoc analysis of data indicated that under learner control, students who exhibited relatively high levels of prior background knowledge were significantly more capable of making accurate self-estimates of understanding and appropriate uses of control options, and thus of utilizing effective learner strategies, than were those with low levels of prior background knowledge: $F(1,29)=4.219$, $p<.05$ for accuracy of self-estimates of understanding; $F(1,29)=24.543$, $p<.001$ for appropriateness of control use; and $F(1,29)=15.400$, $p<.001$ for effectiveness of learner strategy. Students with low levels of prior background knowledge tended to estimate their own understanding inaccurately and to use the control options too infrequently.

These findings on individual differences in prior background knowledge indicate that students' prior knowledge of concepts relevant to those being taught is important in helping them to understand the underlying structure of a given instruction, to develop effective learning strategies, and thus to enhance metacognitive and cognitive knowledge and skills. As Snow (1980) contended, learner controlled instruction may not overcome the fact that "individual characteristics not under control of the individual will determine to a significant extent what and how much that individual will learn in a given instructional setting" (p.1).

Overall, students in the learner control group demonstrated significantly better performance than those in the program control group in the assessments of metacognitive and cognitive effects, regardless of their individual differences in prior background knowledge. This finding contradicts those of some previous research on learner control (Goetzfried and Hannafin, 1985; Hannafin, 1984; Kinzie, 1988; Krendle and Liberman, 1988; Steinberg, 1977), which supported the Achievement-Treatment Interaction theory (Tobias, 1976); for students with a low level of prior knowledge, learner control is less effective than program control, and for those with a high level of prior knowledge, learner control is more effective than (or comparable to) program control. However, although it was theoretically assumed that learner control as an instructional design strategy would serve identified individual differences among learners better than would program control in accomplishing the goal of instruction, the inequity among learners resulting from individual differences was preserved by learner control in the same way as by program control.

The magnitude of individual differences typically confronting instructional designers and the evidence of the relationship between such differences and achievement explain why only a few students in most classrooms fall into the category of "success". The majority of students need some type of support, and thus a variety of supporting strategies must be investigated and integrated into the design of instruction. This applies to the design of learner controlled instruction as well. Further investigation is needed of supportive factors that can take over some of the burdens that learners may have and that may interact with a variety of characteristics of individual differences.

The overall results of this study indicate that as an instructional design strategy, learner control can provide learners with opportunities to improve metacognitive skills and knowledge, while attempting to teach cognitive skills and knowledge of specific subject matter content in a more effective way than can program control. These results were apparent across learners' prior background knowledge of relevant subject matter. However, it must be recognized that this study was conducted in introductory instruction, with a relatively limited application of control options, for a short time period. Considering the potentially rich learning environment that learner control can create, further research efforts should be devoted to the development of instructional models of learner control from which all learners can benefit in enhancing both metacognitive and cognitive skills and knowledge.

REFERENCES

Bonner, J. (1988, Spring). Implications of Cognitive Theory for Instructional Design: Revisited. ECTJ, 36(1), 3-14.

Brown, A. (1980). Metacognitive Development and Reading. In R. J. Spiro et al. (Eds.), Theoretical Issues in Reading Comprehension. Hillsdale, NJ: Lawrence Erlbaum Associates.

Carrier, C. (1984, February). Do Learners Make Good Choices? Instructional Innovator. 15-17, 48.

Chipman, S. F., & Segal, J. W. (1985). Higher Cognitive Goals for Education: An Introduction. In J. W. Segal, S. F. Chipman, & R. Glaser (Eds.), Thinking and Learning Skills. Hillsdale, NJ: Lawrence Erlbaum Associates.

Clememts, D. H., & Nastasi, B. K. (1988, August). Dynamic Approach to Measurement of Children's Metacomponential Functioning. Paper presented at the 96th annual convention of the American Psychological Association. Atlanta, GA.

Cohen, J. (1969). Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press.

Finn, J. D. (1974). A General Model for Multivariate Analysis. New York, NY: Holt, Rinchart and Winston.

Gagné, R. M., & Glaser, R. (1987). Foundations in Learning Research. In R. M. Gagné (Ed.), Instructional Technology: Foundations. Hillsdale, NJ: Lawrence Erlbaum Associates.

Garhart, G., & Hannafin, M. (1986, Summer). The Accuracy of Cognitive Monitoring During Computer-Based Instruction. Journal of Computer-Based Instruction, 13(3), 88-93.

Gay, G. (1986). Interaction of Learner Control and Prior Understanding in Computer-Assisted Video Instruction. Journal of Educational Psychology, 78(3), 225-227.

Goetzfried, L. & Hannafin, M. J. (1985, Summer). The Effects of the Locus of CAI Control Strategies on the Learning of Mathematics Rules. American Educational Research Journal. 22(2), 273-278.

Hannafin, M. J. (1984). Guidelines for Using Locus of Instructional Control in the Design of Computer-Assisted Instruction. Journal of Instructional Development, 7(3), 6-10.

Kinzie, M. B. (1988, April). Learner Control, Continuing Motivation, and Achievement. Paper presented at the annual meeting of American Educational Research Association. New Orleans.

Krendle, K. A., & Liberman, B. A. (1988). Computers and Learning: A Review of Recent Research. Journal of Educational Computing Research, 4(4), 367-389.

Lepper, M. R., & Malone, T. W. (1987). Intrinsic Motivation and Instructional Effectiveness in Computer-Based Education. In R. E. Snow, & M. J. Farr (Eds.), Aptitudes, Learning, and Instruction: VOL.3, Hillsdale, NJ: Lawrence Erlbaum Associates.

Markman, E. M. (1977). Realizing that You Don't Understand: A Preliminary Investigation. Child Development, 48, 986-992.

Merrill, M. D. (1984). What is Learner Control? In R. K. Bass, & C. R. Dills (Eds.). Instructional Development: The State of The Art II. Dubuque, Iowa: Kendall/Hunt Publishing Company.

Perkins, D. N., & Salomon, G. (1989). Are Cognitive Skills Context-Bound? Educational Researcher, 18(1), 16-25.

Rao, C. R. (1973). Linear Statistical Influence and Its Applications: Second Edition, New York, NY: John Wiley and Sons.

Reigeluth, C., & Stein, F. S. (1983). The Elaboration Theory of Instruction. In C. M. Reigeluth (Ed.). Instructional-Design Theories and Models: An Overview of Their Current Status. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schmitt, M. C., & Newby, T. J. (1986). Metacognition: Relevance to Instructional Design. Journal of Instructional Development, 9(4), 29-33.

Snow, R. E. (1980). Aptitude, Learner Control, and Adaptive Instruction. Educational Psychology. 15(3), 151-158.

Steinberg, E. R. (1977, February). Review of Student Control in Computer-Assisted Instruction. Journal of Computer-Based Instruction, 3(3), 84-90.

Tennyson, R. D. (1980). Instructional Control Strategies and Content Structure as Design Variables in Concept Acquisition Using Computer Based Instruction. Journal of Educational Psychology, 72(4), 525-532.

Tennyson, R. D., & Buttrey, T. (1980, Fall). Advisement and Management Strategies as Design Variables in Computer-Assisted Instruction. ECTJ, 28(3), 169-176.

Tobias, S. (1976, Winter). Achievement Treatment Interaction. Review of Educational Research, 46(1), 61-74.

Waldrop, P. B. (1984, September). Behavioral Reinforcement Strategies for Computer-Assisted Instruction: Programming for Success. Educational Technology, 38-41.

APPENDIX A
Tables

Table 1

Means and Standard Deviations of Dependent Variables

Variable	Maximum Score	L Mean	L SD	P Mean	P SD	ALL Mean	ALL SD
utilization of monitoring	24	14.65	3.30	11.42	3.30	13.03	3.66
correctness of monitoring	24	12.77	2.54	9.39	4.15	11.08	3.81
knowledge acquisition	40	25.65	6.70	22.06	8.17	23.85	7.63
knowledge application	70	54.97	14.70	45.65	19.78	50.31	17.91

Note. L = learner control group
P = program control group

Table 2

Results of t Test on Dependent Variables between the Learner Control and Program Control Groups

Variable	D.F.	t Value	One-tail Prob.
utilization of monitoring	60	3.84	.000***
correctness of monitoring	60	3.88	.000***
knowledge acquisition	60	1.89	.032*
knowledge application	60	2.11	.020*

*.01<p<.05. ***p<=.001.

Table 3

Descriptive Data on Accuracy of Self-estimates of Understanding

mean	median	SD
2.503	2.400	.563

Histogram Frequency

```
Count   Midpoint
  0       1.00  |
  1       1.20  |*****
  0       1.40  |
  1       1.60  |*****
  2       1.80  |**********
  3       2.00  |***************
  1       2.20  |*****
 10       2.40  |**************************************************
  3       2.60  |***************
  4       2.80  |********************
  2       3.00  |**********
  1       3.20  |*****
  1       3.40  |*****
  1       3.60  |*****
  1       3.80  |*****
  0       4.00  |
                I....+....I....+....I....+....I....+....I....+....I
                0         2         4         6         8        10
```

Table 4

Descriptive Data on Appropriateness of Control Use

mean	median	SD
2.348	2.400	.611

Histogram Frequency

```
Count   Midpoint
  2       1.00  |**********
  0       1.20  |
  1       1.40  |*****
  2       1.60  |**********
  0       1.80  |
  4       2.00  |********************
  4       2.20  |********************
  7       2.40  |***********************************
  3       2.60  |***************
  1       2.80  |*****
  3       3.00  |***************
  3       3.20  |***************
  1       3.40  |*****
  0       3.60  |
  0       3.80  |
  0       4.00  |
                I....+....I....+....I....+....I....+....I....+....I
                0         2         4         6         8        10
```

Table 5

Descriptive Data on Effectiveness of Learner Strategy

mean	median	SD
2.426	2.400	.522

Histogram Frequency

```
Count   Midpoint
  0       1.00 |
  1       1.10 |**********
  0       1.20 |
  0       1.30 |
  0       1.40 |
  1       1.50 |**********
  0       1.60 |
  1       1.70 |**********
  0       1.80 |
  1       1.90 |**********
  1       2.00 |**********
  3       2.10 |******************************
  3       2.20 |******************************
  1       2.30 |**********
  5       2.40 |**************************************************
  2       2.50 |********************
  3       2.60 |******************************
  4       2.70 |****************************************
  1       2.80 |**********
  0       2.90 |
  0       3.00 |
  0       3.10 |
  2       3.20 |********************
  0       3.30 |
  1       3.40 |**********
  0       3.50 |
  1       3.60 |**********
  0       3.70 |
  0       3.80 |
  0       3.90 |
  0       4.00 |
               I....+....I....+....I....+....I....+....I....+....I
               0         1         2         3         4         5
```

Table 6

Means and Standard Deviations of Dependent Variables

Variable	Maximum Score	E-L Mean	E-L SD	IE-L Mean	IE-L SD	P Mean	P SD
utilization of monitoring	24	14.94	3.80	14.33	2.77	11.42	3.30
correctness of monitoring	24	13.25	2.65	12.27	2.40	9.39	4.15
knowledge acquisition	40	27.75	6.26	23.40	6.61	22.06	8.17
knowledge application	70	58.44	9.78	51.27	18.22	45.65	19.78

Note. E-L = effective learner control subgroup
IE-L = ineffective learner control subgroup
P = program control group

Table 7

Results of ANOVA on Dependent Variables among the Two Learner Control Subgroups and the Program Control Group

variable	source of variation	D.F.	mean square	F Ratio	F Prob.
utilization of monitoring	between group	2	82.058	7.428	.001***
	within group	59	11.048		
correctness of monitoring	between group	2	92.654	7.795	.001***
	within group	59	11.886		
knowledge acquisition	between group	2	172.611	3.180	.049*
	within group	59	54.279		
knowledge application	between group	2	872.605	2.889	.064
	within group	59	302.033		

*.01<p<.05. ***p<=.001.

Table 8

Results of Fisher's Least Significant Difference Test

utilization of metacognitive monitoring

	G1	G2	G3
G1			
G2			
G3	*	*	

correctness of metacognitive monitoring

	G1	G2	G3
G1			
G2			
G3	*	*	

knowledge acquisition

	G1	G2	G3
G1			
G2			
G3	*		

Note. G1 = effective learner control subgroup
G2 = ineffective learner control subgroup
G3 = program control group

Table 9

Results of Regression of Dependent Variables on Pretest Scores

group	r^2	b	statistical value
utilization of metacognitive monitoring			
learner control	.0002	.020	F(1,29) = .005
program control	.0086	.116	F(1,29) = .252
correctness of metacognitive monitoring			
learner control	.0217	.176	F(1,29) = .642
program control	.1540	.615	F(1,29) = 5.281^{*}
knowledge acquisition			
learner control	.1457	1.205	F(1,29) = 4.946^{*}
program control	.2954	1.679	F(1,29) = 12.158^{**}
knowledge application			
learner control	.0218	1.022	F(1,29) = .645
program control	.1823	3.192	F(1,29) = 6.464^{*}

*.01<p<.05. **.001<p<.01.

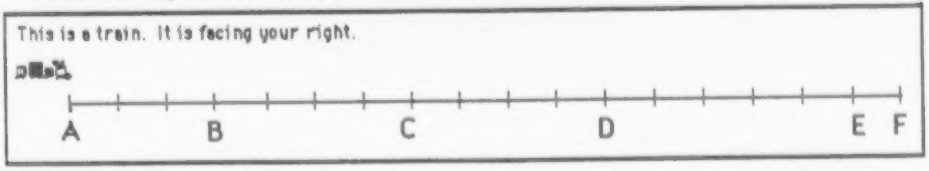
APPENDIX B

Pretest

Your name: ____________________

1] This railroad has been marked in units called blocks.
Each block is ├──┤ long.

A train runs on this railroad.
Use the map of the railroad to help answer the questions below.


This is a train. It is facing your right.

ONE: If the train started at the station B and went to the station E, how many blocks did it move?

Answer: ______ blocks

TWO: The train started at the station C and moved forward 4 blocks. Where did it stop?

Answer: at the ______ station

THREE: The train moved backward 5 blocks. If it ended at the D station, where did it start?

Answer: at the ______ station

FOUR: The train started at the station F. When it left there, it moved backward 5 blocks. How much farther did it have to go to get to the station C?

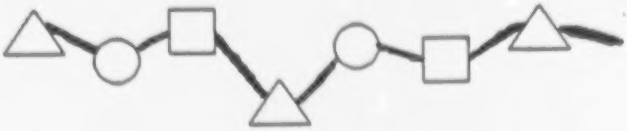
Answer: ______ blocks farther

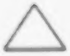
FIVE: The train moved forward 4 blocks and then backward 11 blocks. If it ended at the A station, where did it start?

Answer: at the ______ station

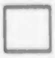
2] Find the pattern and select the shape that comes next.

ONE:

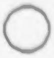




A



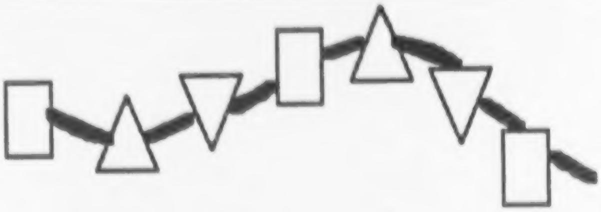
B

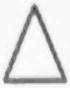


C

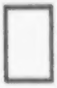
Answer: _____

TWO:

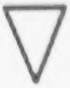




A



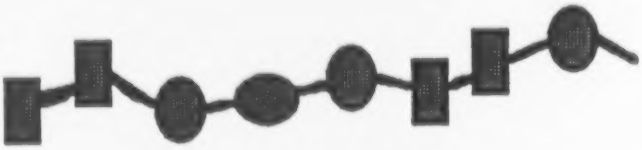
B

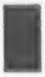


C

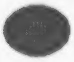
Answer: _____

THREE:

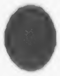




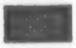
A



B



C

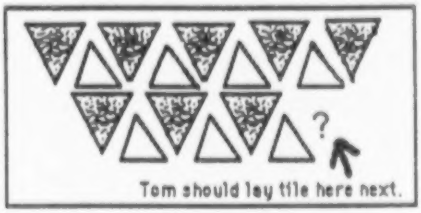


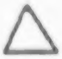
D

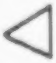
Answer: _____

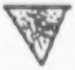
3] Tom is helping his father lay tile in their kitchen.
Which tile, without turning, should Tom use next?
Select the answer among the given choices.

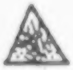
ONE:


?
Tom should lay tile here next.


A

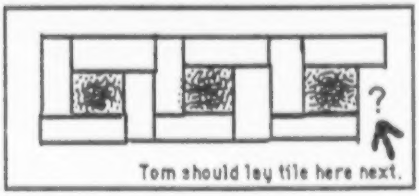

B

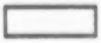

C

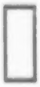

D

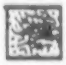
Answer: ________

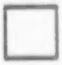
TWO:


?
Tom should lay tile here next.


A

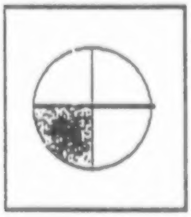

B

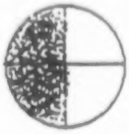

C

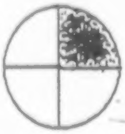

D

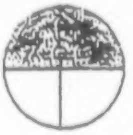
Answer: ________

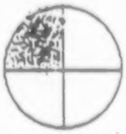
4] If you turn this figure a half turn to the right,
what will the figure look like after the turn?
Select the answer among the given choices.




A

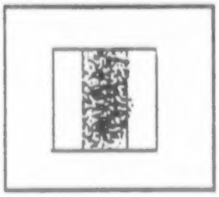

B

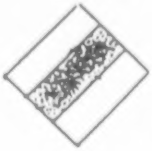

C

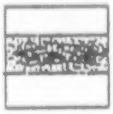

D

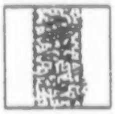
Answer: ________

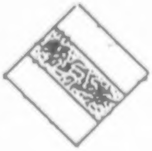
5] If you turn this figure a quarter turn to the left,
what will the figure look like after the turn?
Select the answer among the given choices.




A

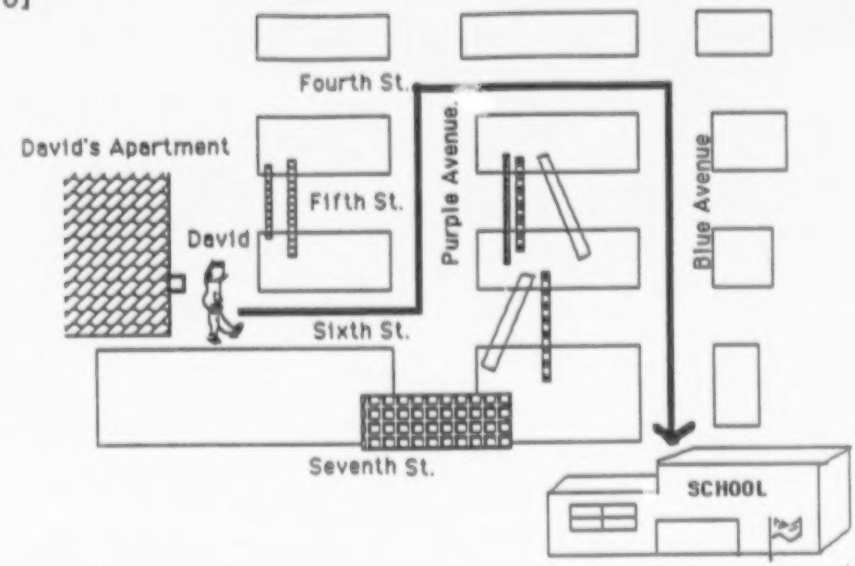

B


C


D

Answer: ________

6]


David's Apartment
David
Fourth St.
Fifth St.
Sixth St.
Seventh St.
Purple Avenue.
Blue Avenue
SCHOOL

Suppose you have to call David to give him directions to get from his apartment house to school.

You have the map, but David doesn't.

As you can see on the map, some roads are closed, so David should follow the path marked on the map.

Answer: Walk along Sixth Street until you come to Purple Avenue, and then turn to the LEFT RIGHT. (Circle one.)

Walk along Purple Avenue until you come to Fourth Street, and then turn to the LEFT RIGHT. (Circle one.)

Walk along Fourth Street until you get to Blue Avenue, and then turn to the LEFT RIGHT. (Circle one.)

Walk along Blue Avenue until you come to the school.

APPENDIX C
Interview on Metacognitive Monitoring

Question #1

Every classroom in Mary's school will have 2 girls for every boy. Mary found out that there will be 8 girls in her class, so she figured there would be 16 boys. Her sister's class will have 10 girls. How many boys do you think will be in her sister's class?

Prompts:

1. Do you have to watch out for mistakes when you do this problem?
2. Is there something in the problem that could trick you if you were not careful?
3. Is Mary right when she figures that there will be 16 boys in her class?
4. Will a classroom with 2 girls for every boy have 2 times as many boys? Is that a mistake? Or half as many boys?
5. Be careful not to make a mistake. If there were 2 girls, and there are 2 girls for every boy, then there would be 1 boy. Half as many boys. If there were 4 girls, then there would be 2 boys. Half as many boys. (Continue if child does not understand until the 8/4 problem of Mary's is solved.)

Question #2

A railroad company discovered that when trains get into an accident, the people in the last car are hurt the most. The last car is the most dangerous. So Sue told them to take off the last car. They did. The train was in an accident. What do you think happened to the people?

Prompts:

1. Do you have to watch out for mistakes when you do this problem?
2. Is there something in the problem that could trick you if you were not careful?
3. Is Sue right when she figures that, if you take off the last car, less people will get hurt?
4. What will happen when you take off the last car? Was there a mistake? Is any car the most dangerous then?
5. Don't make a mistake. If you take off the last car, what would be at the end of the train? Is that car now the last car? Would that car be the most dangerous?

Question #3

When Albert was 6 years old, his sister was 3 times as old as he. Now he is 10 years old and he figures that his sister is 30 years old. How old do you think his sister will be when Albert is 12 years old?

Prompts:

1. Do you have to watch out for mistakes when you do this problem?
2. Is there something in the problem that could trick you if you were not careful?
3. Is Albert right when he figures that when he is 10 years old his sister is 30 years old?
4. Will his sister always be 3 times as old as Albert? Is that a mistake? Should you multiply or add years?
5. Don't make a mistake. When Albert was 6, his sister was 18. 12 years older. What would his sister be 1 year later, when Albert was 7? (19) 12 years older. One year later? (Continue if child does not understand, until solve 10/22 problem of Albert's.)

Question #4

At our school the track is in a circle with six flags spaced equally around it. Yesterday, Frank started at the first flag and raced around the circle. He took 2 minutes to get to the third flag. He thought it would take him 2 more minutes to get to the sixth flag if he kept on at the same rate. How long did it take him to get all of the way around the track?

prompts:

1. Do you have to watch out for mistakes when you do this problem?
2. Is there something in the problem that could trick you if you were not careful?
3. Is Frank right when he figures that it would take him 2 more minutes to get to the sixth flag?
4. Is it the same distance from the first to the third as from the third to the sixth flag? Is that a mistake? Where does he start? How many flags is it to the third flag starting at the first flag?
5. Don't make a mistake. Pretend Frank starts at the first flag, then goes to the second, then the third flag. That takes him 2 minutes. How long will it take him to get to the fourth, then the fifth flag? (The sixth?)

APPENDIX D
Evaluation of Effectiveness of Learner Strategy

Accuracy of Self-Estimates of Understanding

If SELF-EST = 1 & SCORE = 0, 5 or 10, then ACCURACY = 4.
If SELF-EST = 1 & SCORE = 15 or 20, then ACCURACY = 3.
If SELF-EST = 1 & SCORE = 25 or 30, then ACCURACY = 2.
If SELF-EST = 1 & SCORE = 35 or 40, then ACCURACY = 1.

If SELF-EST = 2 & SCORE = 0, 5 or 10, then ACCURACY = 3.
If SELF-EST = 2 & SCORE = 15 or 20, then ACCURACY = 4.
If SELF-EST = 2 & SCORE = 25 or 30, then ACCURACY = 2.
If SELF-EST = 2 & SCORE = 35 or 40, then ACCURACY = 1.

If SELF-EST = 3 & SCORE = 0, 5 or 10, then ACCURACY = 1.
If SELF-EST = 3 & SCORE = 15 or 20, then ACCURACY = 2.
If SELF-EST = 3 & SCORE = 25 or 30, then ACCURACY = 4.
If SELF-EST = 3 & SCORE = 35 or 40, then ACCURACY = 3.

If SELF-EST = 4 & SCORE = 0, 5 or 10, then ACCURACY = 1.
If SELF-EST = 4 & SCORE = 15 or 10, then ACCURACY = 2.
If SELF-EST = 4 & SCORE = 25 or 30, then ACCURACY = 3.
If SELF-EST = 4 & SCORE = 35 or 40, then ACCURACY = 4.

Note. SELF-EST = self-estimates of understanding
SCORE = score in a practice session in a lesson
ACCURACY = accuracy of self-estimates of understanding

Appropriateness of the Use of Control Options

If NOWRONG - NOREVIEW <= 0, then APPROPRIATENESS = 4.
If NOWRONG - NOREVIEW = 1 or 2, then APPROPRIATENESS = 3.
If NOWRONG - NOREVIEW = 3 or 4, then APPROPRIATENESS = 2.
If NOWRONG - NOREVIEW = 5 or 6, then APPROPRIATENESS = 1.
If NOWRONG - NOREVIEW = 7 or 8, then APPROPRIATENESS = 0.

Note. NOWRONG = the number of wrong answers in a practice session in a lesson (NOWRONG = 8-SCORE/5)
NOREVIEW = the number of review options used
APPROPRIATENESS = appropriateness of the use of control options

Effectiveness of Learner Strategy

EFFECTIVENESS = (ACCURACY + APPROPRIATENESS) / 2

Note. EFFECTIVENESS = effectiveness of learner strategy
ACCURACY = accuracy of self-estimates of understanding
APPROPRIATENESS = appropriateness of the use of control options

Title:

Implementation levels of a Videodisc-based Mathematics Program and Achievement

Authors:

William H. Lowry
Ron Thorkildsen

Abstract

We conducted systematic observations of the implementation of a videodisc-based, direct instruction program in fractions.* The nine upper-elementary classrooms involved were spread across three school districts. Implementation/utilization of the program varied across classrooms, as did achievement. Based on the observations we classified three classrooms as high-implementation classrooms, three as average-implementation, and three as low-implemention. At the conclusion of the field study, there was a 28-percentage-point difference between the high-implementing-classrooms mean gain score and that of the low-implementing classrooms. When covariance-adjusted postest means were considered, the difference was about 20 percentage points.

---

* This study was completed as a part of a grant funded by the United States Department of Education, project number 84.180R2--Technology, Educational Media & Materials - Assessing the Social and Cultural Impact of Group-Based, Videodisc Math Instruction in Mainstreamed Classrooms. Its parent study was primarily concerned with teacher-student interactions and achievement outcomes for special education students in mainstreamed classrooms. The present study added the investigation of relationships between program implementation and achievement among regular education students. The studies were conducted concurrently, in the same classrooms. For a more complete description of the research conditions, see Lowry (1989).

BACKGROUND

Many researchers of instructor-controlled instructional technologies try to ensure that the program is implemented as close to design specifications as possible, consistent with Mark's (1983) suggestions (e.g., Kelly, 1986). By doing so, the researchers can verify the treatment variable and so can argue that differences in the dependent variable are attributable to the program (see Shaver, 1983).

Media have been used in an attempt to "standardize" the delivery of instruction (Kemp and Smellie, 1989, p. 3) and so to reduce variations in implementation. But apparently, even when a well-defined, media-based program is put into the classroom and monitored, variations in implementation are observed. Hasselbring, Sherwood, and Bransford (1986) reported that they informally observed such variations during their study of the Mastering Fractions vodeodisc program (Systems Impact, 1986).

In their study Hasselbring et al. (1986) were able to control the implementation of the program to some degree. Larger variations in program implementation are apt to occur when instructors use a program in their day-to-day work, outside the field-test environment (Shaver, 1964; Heinich, 1984). A number of authors have suggested that differences between what an instructional designer recommends and what is implemented in the classroom may have a significant effect on student achievement (e.g., Stake, 1967; Hall and Loucks, 1977).

It is important that program implementation, for pre-designed instructional programs, be gathered systematically (Shaver, 1983). Findings of such research can be used to evaluate the specific program. But information regarding implementation versus the learning outcomes of a program can also be used to evaluate instructional design/development in general. If an instructional design, based on theoretical and empirical foundations, is manifested in an instructional product; the product is utilized "well" in the classroom; and the results are positive; then one of the assumptions regarding instructional design is supported . That assumption is that "... systematically designed instruction can greatly affect individual human development" (Gagne, Briggs, & Wager, 1988, p. 5; see also Dick & Carey, 1990).

Mastering Fractions was designed to be consistent with the Direct Instruction Model (Engelmann and Carnine, 1982) and with the findings of the "effective teaching" literature (see e.g., Engelmann and Carnine, 1982; Rosenshine and Stevens, 1986; Hofmeister & Lubke, 1990). Further, some of the features of the effective teacher are "captured" in the videodisc medium (Hofmeister, personal communication). Therefore, the program was considered an example of a well-grounded and well-defined intervention, based on a systems approach to instructional design and development.

PURPOSE OF THE RESEARCH

To extend the work of Hasselbring et al. (1986), we measured program implementation level as the Mastering Fractions program was being used at three sites. The question specifically addressed by this study was as follows: Does the level of student achievement differ across levels of implementation of a mediated program that is based on a well-defined instructional model?

METHOD

First, to check the overall effect of the Mastering Fractions program, we used a quasi-experimental, nonequivalent control group design to compare the postest achievement of two groups of upper elementary students (N = 337; 171 treatment, 166 control). The treatment group received instruction in fractions via the teacher-directed, videodisc-based, Mastering Fractions program. The control group students received their normal grade-four or grade-five mathematics program, which did not include common fractions. This part of the study provided data regarding treatment vs. no-treatment conditions. The mean, treatment-group, covariance-adjusted, posttest score on a criterion-referenced achievement test of fractions skills and concepts was 6.0 standard deviations above that of the control group (the effect size = 5.9 for raw gain scores). These effect sizes were determined using the student as the unit of analysis. We concluded that the program did provide solid instruction in specific fractions skills and concepts.

In addition to the comparison between the treatment and control groups, we collected data on differences in achievement within the treatment group (n = 171) by level of program implementation. This addressed the research question posed in the present study.

Subjects

The facilitators involved in this study were nine grade-five teachers at three sites. Each had a minimum of five years of teaching experience.

The students were grade-five, public school students who had not received instruction in fractions. Their grade equivalent scores (based on the Iowa Test of Basic Skills or the Comprehensive Test of Basic Skills) are comparable to those presented in the technical manuals for the instruments (see Lowry, 1989).

Instrumentation

The two variables measured for this study were criterion-referenced student achievement and teacher implementation of the program. The former was measured using a paper-and-pencil test; the latter was considered using a combination of measures.

Achievement measure

We used a criterion-referenced test, developed for previous research with Mastering Fractions, as a measure of fractions achievement. The test contains 57 items covering fractions skills and concepts. Since it is criterion referenced, there were low scores and low variability among scores at pretest. The test-retest correlation for control-group students is +0.67.

Level of implementation of the Mastering Fractions program

Shaver (1983) suggests the use of multiple measures of implementation, including systematic observation. He defines systematic observation as "... the unitization of behavior and the classification of the units into previously defined categories,..." (p. 4). We used three measures of implementation: a systematic-observation instrument, a self-report teacher questionnaire regarding program modification, and classroom seating charts.

Observational instrument. Although implementation observation instruments are available, Shaver (1983) points out that these instruments "... may not reflect validly the critical dimensions of the teaching methods to be investigated" (p. 4). The developers of Mastering Fractions provide specific suggestions for implementing the program. These suggestions are presented in the teacher's manual. We combined the suggestions from the manual with two other sources: (1) information from informal observations made by experienced observers from earlier studies of the program and (2) analyses of videotapes of teachers who had used the program successfully in the past. Once the behaviors were listed, we categorized them according to a media utilization model (Kemp & Smellie, 1989) and the effective teaching literature (see Hofmeister & Lubke, 1990). The result was a dual-purpose listing. We used the list as a rating instrument for systematic observation and as a set of suggestions for teachers implementing the program. The suggestions sheet is presented in Appendix A; the rating sheet appears in Appendix B.

During the course of the Mastering Fractions program in each of the 9 treatment-group classrooms in the fall of 1988, observers for the parent project were asked to record their perceptions of the degree of implementation of each area of the utilization model (see Appendix B). The sum of the rating points (on a five-point scale; 5 as high), for all of the categories in the utilization model, was taken as an implementation score on this instrument.

At least two observers would visit a single class together, once per week, to check interobserver agreement. One investigator also visited each classroom where the videodisc program was being used, to check interobserver agreement. Percentage of exact interobserver agreement was 75.6; agreement within one step on the rating scale was 96.7%. The latter percentage is provided, since it is quite conceivable that, on a five-point scale representing a continuous variable, one observer might perceive a behavior as just less than say 2.5 out of 5 (and so would record a score of 2) while another observer might see the same behavior as slightly more than 2.5 (and so would record it as a score of 3).

Teacher questionnaire. After each of the classes had completed the Mastering Fractions program, site coordinators interviewed the respective teachers with a structured questionnaire as a self report of modifications to the program. A copy of the questionnaire appears in Appendix C. The questionnaire provided a second indication of how teachers modified the recommended delivery practices.

Seating charts. The teacher's guide also includes suggestions for seating patterns, to allow the teacher to provide needed help to students who may be experiencing difficulties with the program. The teachers provided seating charts to the observers, so that the observers could identify students, by name, for the parent-study observations. We used the seating charts to determine the degree to which teachers were able to arrange the class seating in accord with the suggestions.

We then compiled the data to get an overall-implementation consistency rating for each teacher, based on the questionnaire, the implementation observations, and the seating charts.

The Instructional Programs

Teacher in-service instruction

Treatment-group teachers were given a half-day workshop on the use of the program, and were invited to call the researchers with their questions or comments. This procedure is consistent with the typical in-service training for the program.

During each of the teacher-training sessions we provided copies of the suggestions sheet (Appendix A) and discussed each item on the list. We demonstrated the set-up and use of the equipment. We addressed each of the program utilization techniques by showing and discussing a videotape of a model teacher using the program. We encouraged teachers to follow the suggestions as closely as possible.

The Mastering Fractions program

Rosenshine and Stevens (1986) provide a list of teaching processes in their operational definition of the daily functions of direct instruction:

1. Review, check the previous day's work (and reteach, if necessary)
2. Present new content/skills

3. Guide student practice (and check for understanding)
4. Provide feedback and correctives (and reteach, if necessary)
5. Provide independent student practice
6. Provide weekly and monthly reviews.

But rote following of these instructional functions does not guarantee that students will learn. Rosenshine and Stevens (1986) noted that the success of these techniques depends on specific ways of performing each of them. For example, content or skills should be presented in a clear and organized manner (item 2 above). Haphazard presentation does not benefit the learner (see also Bandura, 1982).

Direct forms of instruction have been observed by researchers studying effective-teaching in practice (e.g, Good, 1979). But the Direct Instruction Model (Engelmann, 1980) has evolved from a theoretical/empirical base. For a detailed description of the model see Engelmann and Carnine (1982) (see also Becker & Carnine, 1980).

The Mastering Fractions program was developed in cooperation with Engelmann and Carnine. It was designed to be presented to a heterogeneous class of students. The program requires that a television or monitor, with a screen large enough to be read in all parts of the classroom, be placed at the front of the class. All students are seated facing the screen. The audio system must be clear and loud enough to be heard by all students. The program requires the Mastering Fractions videodiscs, a level-one videodisc player, consumable student worksheets, extra paper and pencils, and the teacher handbook (which includes answer keys for the seat work and tests).

Mastering Fractions consists of 35 lessons. Each lesson is designed to be completed in approximately 40 minutes. Most presentation/practice lessons begin with a quiz of the previous lesson's learning, and each fifth lesson is a mastery test. Both quizzes and mastery tests are followed by diagnostic/remedial suggestions based on specific skill weaknesses manifested by the students. Instruction is provided on the disc at three progressive levels--oral response, written practice with component skills and concepts, and written practice with articulated skills and problem-solving strategies (for more detail see Systems Impact, 1986).

The developers of the Direct Instruction Model have provided specific guidelines for the presentation and practice of concept attainment and skill performance. That is, general definitions of direct forms of instruction address instructional management or a superstructure for instruction, while the Direct Instruction Model also provides instructional design rules and message design guidelines. The videodisc-based portion of the Mastering Fractions program was designed using the D. I. Model, and the paper-based materials and recommended teaching procedures are designed to support that model. However, most important to the present study, while the program is media based, it is teacher controlled. The designers purposely gave the teacher control over portions of the program to help integrate Mastering Fractions into the school context.

Analysis

Each teacher did modify the delivery of the Mastering Fractions program to some degree. When the treatment-group classes had finished the program, the observations were averaged for each teacher on a per-observation-period basis. The post-study questionnaire data and seating chart data were also included. The seating charts showed that none of the teachers used the suggested seating arrangement consistently. Scores on the post-study questionnaire were consistent with the observational data. We then classified the classes as high-implementing, average-implementing, and low-implementing--three classes in each subgroup.

Adequate analysis was made by simple inspection of the data (Shaver, 1983) (see RESULTS AND DISCUSSION). We also ran an analysis of covariance, using pretest means on a standardized achievement measure and pretest means on the criterion-referenced test as covariates.

RESULTS AND DISCUSSION

Achievement and Program Implementation

Table 1 shows mean percentage scores and gain scores on the fractions criterion-referenced test. The treatment-group means are broken down by the level of program implemenation. The the no-treatment control group means are also included for your information. The overall mean of the treatment group is also included.

Table 1

Percentage Correct Means on the Criterion-Referenced Achievement Test by Program Implementation Level for the Treatment Group

Implementation Level	Pretest		Posttest		Gain	n-size
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	
High	6.6	5.0	84.9	10.0	+78.3	55
Average	11.5	15.3	79.7	14.6	+68.2	54
Low	8.6	11.2	58.9	24.8	+50.3	62
Control	10.6	10.7	9.2	10.6	-01.4	166
Overall within the treatment group	8.9	11.4	73.8	21.3	+64.9	171

As would be expected the control group scores show essentially no change on the fractions criterion-referenced test from from pre- to posttest. The treatment group means, however, show large gains (64.9 percentage points on the average). As noted earlier, overall, the results of the criterion-referenced test provide evidence that students who study the Mastering Fractions program do learn specific fractions skills and concepts. The standardized mean difference effect size between the treatment and control groups was +5.9 in favor of the treatment group. The mean scores on the criterion test were also consistent with gains made in previous research on this instructional system (e.g., Hasselbring et al., 1986).

There are also clear differences among the means for high, average, and low levels of implementation. There is a 10 percentage-point difference between the high- and average-implementation mean gain scores, and an 18 percentage-point difference between the average- and the low-implementation mean gain scores.

The analysis of covariance yielded an effect of levels-of-implementation that is statistically significant when the student is used as the unit of analysis ($F = 25.92$; $p = .001$). The eta-squared effect size, based on the adjusted means, is +0.12, which means that 12% of the variance in the criterion-referenced posttest scores is associated with differences in implementation level. Table 2 shows the covariance-adjusted posttest means for the criterion-referenced posttest scores (for details see Lowry, 1989). There remained a 20-percentage point spread between the mean of the high-implementation classes and that of the low-implementation classes. No site dominated any of the implementation levels.

Table 2

Summary of Adjusted Cell Means for the Analysis of Covariance

Level of Implementation	Mean
High	83.6%
Average	76.1%
Low	63.9%

Discussion

The findings of this study are consistent with observations made by Hasselbring et al. (1986). Variations in implementation/utilization did occur with an instructional system that provides for some autonomy on the part of the teacher.

Further, the achievement on a criterion-referenced test appears to be directly related to the level of implementation of the program. This relationship supports an assumption that systematically designed instruction can affect individual human development. In this case the area was achievement enhancement, the goal of a direct-instruction-type program.

The results also support the position that systematic design is necessary but insufficient. If a causal relationship exists between levels of implementation of a well-designed program and achievement, for this or other systems-designed programs, designers must be cautious. If an instructional system is too sensitive to the context in which it is implemented, i.e., too open, its potential effect may be watered-down (as noted by Heinich, 1984).

References

Bandura, A. (1982). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.

Becker, W. C., Carnine, D. W. (1980). Direct instruction: An effective approach to educational intervention with the disadvantaged and low performers. In B. B. Lahey & A. E. Kazdin (Eds.), Advances in clinical child psychology (Vol. 3), pp. 429-473. New York: Plenum.

Dick, W., & Carey, L. (1990). The systematic design of instruction (3rd ed.). New York: Harper Collins.

Engelmann, S. (1980). Direct instruction. Englewood Cliffs, NJ: Educational Technology Publications.

Engelmann, S., & Carnine, D. (1982). Theory of instruction: Principles and applications. New York: Irvington.

Gagne, R. M., Briggs, L. J., & Wager, W. W. (1988). Principles of instructional design (3rd ed.). New York: Holt, Rinehart, and Winston.

Good, T. L. (1979). Teacher effectiveness in the elementary school. Journal of Teacher Education, 30(2), 52-64.

Hall, G. E., & Loucks, S. F. (1977). A developmental model for deterkming whether the treatment is actually implemented. American Educational Research Journal, 14(3), 263-276.

Hasselbring, T., Sherwood, B., & Bransford, J. (1986). An evaluation of the Mastering Fractions level-one instructional videodisc program. Nashville, TN: The Learning Technology Center, George Peabody College of Vanderbilt University.

Heinich, R. (1984). The proper study of instructional technology. Educational Communications and Technology Journal, 32(2), 67-87.

Hofmeister, A., & Lubke, M. (1990). Research into practice: Implementing effective teaching strategies. Boston, MA: Allyn and Bacon.

Kelly, B. F. (1986, August). The relative effectiveness of instuctional design features in two fractions curricula with mildly handicapped remedial high school students. Unpublished doctoral dissertation, University of Oregon, Eugene.

Kemp, J. E., & Smellie, D. C. (1989). Planning, producing, and using instructional media. New York: Harper & Row.

Lowry, W. H. (1989). The effects of a direct instruction program in fractions on academic and mathematics self-concept. Unpublished doctoral dissertation, Utah State University, Logan.

Mark, M. M. (1983). Treatment implementation, statistical power, and internal validity. Evaluation Review, 7(4), 543-549.

Rosenshine, . V., & Stevens, R. (1986). Teaching functions. In M. Whittrock (Ed.) Third handbook of research on teaching. New York: Macmillan.

Shaver, J. P. (1964). The ability of teachers to conform to two styles of teaching. The Journal of Experimental Education, 32, 259-267.

Shaver, J. P. (October, 1983). The verification of independent variables in teaching methods research. Educational Researcher, pp. 3-9.

Stake, R. E. (1967). The countenance of educational evaluation. Teachers College Record, 68(7), 523-540.

Systems Impact, Inc. (1986). Mastering fractions. [Videodiscs and print]. Washington, DC: Author.

SUGGESTIONS FOR THE USE OF "MASTERING FRACTIONS"

9/20/88

Preparing Yourself
- practice setting up the equipment.
- practice using the program before you teach with it.
- look for subtle variations.
- go through the write/check screens and the worksheets.
- practice with the hand controller.
- look for potential problem areas for your students.

Preparing Your Environment
- ensure the equipment is ready.
- ensure the workbook pages are photocopied.
- ensure the lighting level is appropriate.
- ensure there is no glare on the TV screen.
- ensure all students can read the screen and hear the audio track.
- ensure there is adequate space for you to walk around the room.

Preparing the Students
- note the lesson number for today aloud.
- call their attention to the information on the screen (help them focus on the screen)
- ensure students have paper, sharp pencils, etc.
- position students having difficulty near the center of the group.
- stop disruptions.
- introduce support skills.
- support the program verbally and by your actions.

Evaluating Previous Learning
- use the review quiz.
- allow adequate work time for each problem.
- use quiz results to determine remediation.
- use the "1/5 criterion"
- use progress sheets.
- allow student mastery, not a weekly schedule, to determine program pacing.

Presenting and Remediating the Instructional Material
- circulate among the students.
- model/encourage appropriate student verbal response.
- use the scan and step functions to freeze frame or review/reiterate concepts or vocabulary.
- stop the program to clarify or enhance the instruction if students are having difficulty.

Guiding Practice
- circulate to check student progress on their work products.
- diagnose student problems.
- coach students, a step at a time, through new types of problems, using previously learned skills.
- keep the pace moving.
- have alternate work available for high achieving students (they can switch their attention from math to another project and back, as necessary).
- work briefly with individuals, then circulate; review/remediate with the group as necessary.

Appendix B

Implementation Rating Sheet

SUGGESTIONS FOR THE USE OF "MASTERING FRACTIONS"

9/20/88

Preparing Yourself 1 __ 2 __ 3 __ 4 __ 5 __
- practice setting up the equipment.
- practice using the program before you teach with it.
- look for subtle variations.
- go through the write/check screens and the worksheets.
- practice with the hand controller.
- look for potential problem areas for your students.

Preparing Your Environment 1 __ 2 __ 3 __ 4 __ 5 __
- ensure the equipment is ready.
- ensure the workbook pages are photocopied.
- ensure the lighting level is appropriate.
- ensure there is no glare on the TV screen.
- ensure all students can read the screen and hear the audio track.
- ensure there is adequate space for you to walk around the room.

Preparing the Students 1 __ 2 __ 3 __ 4 __ 5 __
- note the lesson number for today aloud.
- call their attention to the information on the screen (help them focus on the screen)
- ensure students have paper, sharp pencils, etc.
- position students having difficulty near the center of the group.
- stop disruptions.
- introduce support skills.
- support the program verbally and by your actions.

Evaluating Previous Learning 1 __ 2 __ 3 __ 4 __ 5 __
- use the review quiz.
- allow adequate work time for each problem.
- use quiz results to determine remediation.
- use the "1/5 criterion"
- use progress sheets.
- allow student mastery, not a weekly schedule, to determine program pacing.

Presenting and Remediating the Instructional Material 1 __ 2 __ 3 __ 4 __ 5 __
- circulate among the students.
- model/encourage appropriate student verbal response.
- use the scan and step functions to freeze frame or review/reiterate concepts or vocabulary.
- stop the program to clarify or enhance the instruction if students are having difficulty.

Guiding Practice 1 __ 2 __ 3 __ 4 __ 5 __
- circulate to check student progress on their work products.
- diagnose student problems.
- coach students, a step at a time, through new types of problems, using previously learned skills.
- keep the pace moving.
- have alternate work available for high achieving students (they can switch their attention from math to another project and back, as necessary).
- work briefly with individuals, then circulate; review/remediate with the group as necessary.

Appendix C

Teacher Questionnaire

Teacher ____

Date of Interview ____________________________________

During your post- Mastering Fractions interview, we collected information regarding your feelings about the program. The program developers are also interested in how you used Mastering Fractions. They would like any information on how you may have modified the program.

1. How many weeks do you normally spend on fractions, at the grade level you're teaching this year?

    ____________________ weeks

2. Was there terminology or content in the Mastering Fractions program which you modified. If so, what specific terms or concepts stand out in your mind?

3. Were there instructional methods used in the program which you modified? Specifically . . . . .

    a. Did you modify the 1/5 criterion for remediation, which appears on the "CSP" screens? If so, how?

    b. Did you require students to "write" or "copy" problems where the program asked students to "copy the problem and work it"? If not, what other approach(es) did you use?

c. Did you require choral responses during the verbal practice presented in the program?

d. Did you use choral responses in place of written responses? In what parts of the program?

4. Did the students use the workbooks as homework or did they use them as independent seatwork? (Please circle one or both - if both apply, then state the approximate percent of time for each.)

5. Did you supplement the videodisc material with other materials? If so, during what parts of the program, or regarding what concepts/strategies?

6. During the Fall session, USU Staff developed "Supplements" intended to help students learn multiple labels for the concepts they had learned during Mastering Fractions; that is, to generalize the concepts they were learning.

Did you use the Supplements? ______________________

If so, do you think your students benefited from their use?

7. Do you think that _Mastering Fractions_ is appropriate for the grade level you teach? ________________

If not, what recommendations would you have for its use?

8. Was there a point in the _Mastering Fractions_ where your lowest students began to experience difficulty? ____________

If so, at about what lesson? ______________________________

Title:

Learner Control versus Computer Control in Instructional Simulation

Authors:

Joseph S. Mattoon
James D. Klein
Richard A. Thurman

Abstract

This experiment was designed to assess the effects of learner control over level of challenge in computer-assisted instruction (CAI) and collect data that will guide the design of future learner control experiments. The CAI included a simulation of an aviation instrument and taught subjects how to locate an aircraft's position in space by interpreting a complex display. Instructional research has not yielded empirical information necessary to develop prescriptions for employing learner control features in this type of CAI. Three treatments that varied the subject's control over level of challenge were compared: learner control, computer control, and learner control with advisement. Results indicated that learner control did affect subjects' time on task on an immediate test, but this effect appeared to be temporary. The change in time on task was likely caused by the way subjects adjusted their challenge level. Subjects who received advisement tended to set challenge levels higher and adjusted challenge more frequently.

LEARNER CONTROL VERSUS COMPUTER CONTROL IN INSTRUCTIONAL SIMULATION

State-of-the-art computer-assisted instruction (CAI) authoring systems enable instructional developers to assign control to learners or use a program to automatically control a variety of instructional components including lesson pace, sequence, feedback, item difficulty, and study strategies. Learner control of these components has been shown to significantly impact the time learners spend on a CAI lesson (Ross & Rakow, 1981), their performance (Avner, Moore, & Smith, 1980), and their continuing motivation to complete CAI tasks (Kinzie & Sullivan, 1989). This indicates that learner vs. computer control may be an important variable in the design of CAI. Despite this, empirical research has not led to useful prescriptions that specify which aspects of CAI should be controlled by the student and which should be controlled by the computer to produce the most effective instruction (Goetzfried & Hannafin, 1985; Klein, 1988). This paper reports the results of research on learner vs. computer control in dynamic simulation-based CAI. The end goal of the research is to advance prescriptions which will help instructional developers in designing effective CAI of this nature.

In certain dynamic tasks that require critical decisions to be made based on information presented in displays, speed can be equally as important as accuracy (Schneider, 1985). For example, pilots must constantly interpret information from different displays, make decisions based on their interpretations and follow through with actions to successfully control their aircraft. Unlike measures of performance in classroom settings (e. g., grading methods for academic tests), speed and accuracy are measured on a continuum with no clear upper or lower bounds. Less error and higher rates of speed indicate progress toward mastery. Such skills often require hours of practice and the actual technical apparatus student pilots are being trained to use is expensive. Also, flaws in performance during the operation of a real aircraft can be dangerous. Therefore, CAI is used to simulate instrument displays and provide a means for student pilots to perfect their skills in advance of training in an aircraft. This type of CAI is referred to as instructional simulation. Little research has been conducted with this type of CAI. Reigeluth and Schwartz (1989) have indicated that more empirical data is needed to develop prescriptions that will lead to effective instructional simulation.

One method of motivating learners to increase their rate of speed and degree of accuracy when performing a dynamic task is to provide higher "levels of challenge" as their performance improves. This is a strategy that has been described as one of the most motivating aspects of computer games (Malone, 1981). Challenge can be increased or decreased by changing the criterion by which learners are given points for good performance. For example, a low challenge level can be defined as follows: 10 points are added to the learner's score for performing a task within 30 seconds, but these points will be subtracted if the learner exceeds this time limit. A high level of challenge for the same task would be produced by reducing the time limit to 15 seconds and increasing the number of points (to be gained or lost) to 100. Instructional simulations can be designed to either allow learners to control level of challenge, or challenge can be controlled by the program as a function of the learner's performance. Klein (1988) states that "...learners may not be the best judges when given control over item difficulty" (p. 25). They tend to choose too high or too low a challenge level for their current level of mastery.

Johansen and Tennyson (1983) argue that learners should be provided with information about their performance during a practice activity. They found that learners achieved more during a CAI lesson when given ongoing information on their degree of mastery of the learning objectives. Tennyson (1980) demonstrated that subjects performed better and completed instructional objectives in less time when using CAI that provided advice and allowed control over instructional components. In general, Tennyson and associates have shown that learner control features alone are not particularly effective, but if CAI provides useful advice throughout the learning activity that is relevant to the criterion task, learner control features can increase the efficiency of the instruction.

Despite the findings that learner control combined with advice can be an effective instructional component of CAI, most CAI research has been restricted to the acquisition of skills involving declarative and conceptual knowledge. The majority of CAI programs deliver content material in the form of verbal information, and students demonstrate mastery of the objectives by choosing answers to multiple-choice questions or by responding to questions with short definitions or descriptions. However, acquiring the skill of making decisions based on information from a dynamic visual display is quite different. Instead of studying a volume of text containing a variety of facts and concepts and attempting to understand and remember the information, mastering skills for interpreting complex instrument displays requires hours of practice on the same basic task. For example, a student pilot must learn to integrate text, digital, symbolic, and spatial information in a single visual display. A decision must then be made based on this information, and the decision is acted on by executing a physical task (e. g., adjust the side-stick controller) to control the aircraft. Such a task must be executed with a high degree of speed and accuracy. Developing skills of this nature involves perceptual, motor, and cognitive ability and may require several hundred practice attempts across a variety of instrument settings. The experiment described below was conducted to examine the effects of learner control over level of challenge with and without advisement. Further, it examines learners' use of control features and their effects on performance and retention of skill for interpreting a complex visual display.

Method

Participants

Subjects were 78 undergraduates enrolled in the same course in the College of Education at a large public university in the Southwestern United States. They participated in the experiment for partial course credit. Performance data was collected from 12 pre-experimental subjects (four per treatment). The remaining 66, 12 males and 52 females, participated in the experiment. This male to female ratio is common for the college.

CAI Program

A CAI program, entitled HUD JR, was developed at AFHRL to conduct the research. HUD JR simulates part of a cockpit display called the Head-Up Display (HUD) and teaches subjects how to interpret its digital and symbolic information. One of the functions of the HUD is to provide radar-generated information on the position of other aircraft. This information is displayed as a set of numbers and symbols that must be interpreted by a pilot to determine location and heading of a target aircraft.

Procedure

The experiment was a two variable mixed design with repeated measures on test occasion. It had three levels of learner control, learner control with advisement (LCA), learner control without advisement (LC), and computer control (CC) by two levels of test occasion, immediate vs. seven day delayed. Three treatment groups were prepared by loading one of three versions of HUD JR on each computer in a microcomputer laboratory. Data from the pre-experimental subjects was used to develop levels of challenge for speed and accuracy. Subjects were randomly assigned in about equal numbers to treatments based on the order they arrived at the computer laboratory. Six groups consisting of six to 20 subjects, equally distributed across the three treatments, completed two experimental sessions. In the first session, subjects completed a short lesson and skill test for using the Macintosh mouse, a lesson on interpreting the HUD, a 75-trial practice exercise for interpreting the HUD, a 25-trial test for locating targets with the HUD, and a questionnaire, all of which were delivered on the Macintosh II computer. The mouse skill test was used to measure subjects' initial ability for moving an icon with the mouse. Subjects were urged to pay close attention to the visual and verbal feedback provided during practice and use it to help them improve their performance. During the lesson, they were frequently reminded that speed and accuracy are equally important. The first session lasted about two hours. Subjects were instructed not to discuss the experimental procedures or the materials among themselves or with their classmates prior to the final session. The second session was held exactly one week later. Subjects spent about 10 minutes completing a second 25-trial test for locating targets with the HUD. After all subjects had completed both sessions, they were debriefed and provided with feedback on their performance. The data for five subjects was eliminated from the analyses, four due to technical problems with the software and one because of failure to complete the delayed test.

Materials

The experimental materials consist of five computer programs that: (1) taught subjects how to use a mouse to manipulate an icon on the CRT; (2) taught the basic knowledge and skills required for interpreting the HUD and locating targets; (3) provided practice and feedback for the target-location task; (4) delivered a posttest for target location; and (5) delivered a questionnaire to record subjects' attitudes and judgments about the learning experience. The programs recorded various behavioral data (e. g., time on task and accuracy of response) which was stored on disk in a standard text format.

The instruction and content material for HUD JR was validated in two ways. First, the basic target-location task was patterned after previous CAI programs developed at AFHRL (Lahey & Boyer, 1990). The earlier programs were completed following a task analysis of the target-location task and interviews with student and instructor pilots. Second, subject matter experts reviewed HUD JR and provided comments that were used to guide revisions prior to the experiment.

Learners are taught to read and interpret three types of information displayed on the HUD as shown in Figure 1: range (distance to the target aircraft displayed as a number from 15 to 40 miles), azimuth (degrees of the target to the left or right of the pilot's heading indicated as an arrow pointing to the right or left and a number from 0 to 60), and aspect (angle from 0 to 359 degrees used to identify heading of the target and displayed as a pointer located on the perimeter of a circle).

Please insert Figure 1 about here.

During practice and test activities, HUD JR simultaneously presents values for each of the three types of information producing a variety of targets, none of which are repeated.

The target-location task requires learners to interpret the information displayed in the HUD and place the "Locator" icon (resembling an aircraft) at the correct position in the "Overhead," an overhead view of the pilot's aircraft and surrounding air space (Figure 1). After positioning the Locator, the subject saw (1) the target (icon showing the position where the subject should have placed the Locator); (2) the "challenge meters;" and (3) the number of points gained or lost for speed and accuracy (Figure 2).

Please insert Figure 2 about here.

The learner evaluates his/her performance by comparing the position of the Locator with the position of the target and by examining the number of points gained or lost. The number of points gained or lost was a function of the learner's speed, accuracy, and current challenge level.

Three experimental treatments were created by developing different versions of the HUD JR program. The LC treatment allowed the learner to set levels of challenge for speed and accuracy by moving the pointers of the challenge meters up or down the scales with the mouse. The LCA treatment gave subjects the same control but also displayed advice (e. g., "set the speed challenge to 35") based on their cumulative performance over the most recent five trials. In the CC treatment the challenge levels were automatically set and learners could not alter challenge for speed or accuracy. The target-location task was the same as in practice for both immediate and delayed tests except that no visual (appearance of the target) or verbal (points) feedback was provided. The subjects responded to bipolar attitude statements about their learning experience and answered questions about their performance after the immediate test by moving a pointer along a linear scale (Figure 3).

Please insert Figure 3 about here.

Criterion Measures

Subjects' performance consisted of their speed and accuracy for locating targets. Additional data was collected to further describe subjects' behavior during the learning activities.

Subjects' response speed for locating a target included the time interval from the onset of the information displayed in the HUD to the time the subject clicked on the Continue button. A subject's accuracy was judged on three dimensions: range error (in centimeters), azimuth error (in degrees), and heading error (in degrees).

Responses to attitude statements and questions were recorded as the position of a pointer along a linear scale. Subjects were asked to estimate aspects of their performance such as time on task and number of targets located before becoming fatigued.

Data Analysis

Multivariate analyses of variance (MANOVA) were conducted on dependent measures of performance and on-task behavior. Additionally, univariate tests were conducted on individual variables. Initial task time for positioning the icon with the mouse was used as a covariate for multivariate and univariate tests when it showed a significant correlation with the dependent variable of interest. The significance level was set at $p<.05$ for all multivariate and univariate tests.

Results

Performance on Immediate and Delayed Tests

Table 1 displays the means for each of the dependent variables by treatment group. Smaller means equate to superior performance since they indicate greater speed or less error.

Please insert Table 1 about here.

The MANOVA did not reveal a significant main effect for learner control treatments ($F(8,108)=0.62$, $p>.05$), but a main effect did occur for test occasion ($F(4,55)=6.02$, $p<.001$). The means in Table 1 reflect this increase in task time and error from the immediate to the delayed test across all four dependent measures.

Univariate analyses revealed an interaction for treatment by test occasion ($F(2,58)=4.95$, $p<.05$). LL subjects performed the target-location task more quickly on the immediate test than the CC or LCA subjects ($M=12.9$ vs. $M=14.1$ and $M=13.7$ seconds) but performed closer to the same rate on the delayed test ($M=14.6$ vs. $M=14.4$ and $M=15.7$). None of the other performance variables showed significant differences in performance as a function of the learner control treatments or test occasion.

On-Task Behavior

Because the difference on task time between the two tests was most apparent for the LC and LCA treatments, a second MANOVA was conducted on the challenge levels and points gained and lost during the practice activity for these two groups. It revealed a significant multivariate effect for treatment ($F(2,38)=4.48$, $p<.05$). The univariate tests showed that the LCA subjects set a higher challenge level for accuracy than the LC subjects ($M=60.2$ vs. $M=41.3$, $F(1,39)=8.91$, $p<.01$) but gained less accuracy points ($M=38.5$ vs. $M=56.7$, $F(1,39)=4.92$, $p<.05$).

A third MANOVA was conducted on the LCA and LC treatment groups for the number of changes subjects made in level of challenge for speed and accuracy. It revealed a significant multivariate effect for treatment ($F(4,36)=5.61$, $p<.001$). Univariate tests indicated that the LCA subjects changed the level of challenge more frequently than did subjects in the LC group: $M=14.1$ vs. $M=6.4$, $F(1,39)=23.4$, $p<.001$ for accuracy challenge and $M=10.0$ vs. $M=5.3$, $F(1,39)=10.71$, $p<.01$ for speed challenge.

To determine to what degree subjects in the LCA group followed the challenge advice delivered every five trials, the probability that subjects adjusted the level as prescribed was computed. This was done for every occurrence of advice where the current level of challenge was at least five units higher or lower than the prescribed level. The probability that advice for speed was followed (to within five units of the prescribed level) was $p=0.75$ and was $p=0.68$ for accuracy.

Responses on Questionnaire

Responses to some of the questions about the learning experience were useful in describing behavior. There was no significant effects for treatments for judgments subjects made on "...number of targets you located before becoming tired of the task" or for the attitude statement, "I enjoyed locating targets with the HUD Simulator." Means for these two items were computed across treatments. The mean for number of targets located before subjects begin to feel tired of the task was 31. The mean for the attitude statement was 0.30 which inferred that subjects tended to enjoy the task. The response range for attitude statements was -1.0 (complete disagreement) to 1.0 (complete agreement). Finally, there was a significant effect ($p<.01$) for treatments for the attitude statement, "I would want the level of challenge for speed controlled by the computer," M=-0.21, 0.20, and -0.23 for LCA, CC, and LC treatments respectively. This inferred that learner control was favorable to subjects in groups that employed learner control, but computer control was favorable to subjects in the computer control treatment. A similar statement for control of accuracy challenge did not reach a level of significance, M=-0.05, 0.11, and -0.19 for LCA, CC, and LC respectively.

Discussion

Because the learner control treatments did not permanently effect performance, it is assumed that learner control did not change subjects' ability to learn the task. Detecting performance effects of learner control for this type of task may require multiple practice sessions. Cronbach and Snow (1977) intimated that 10 or more separate interactive sessions may be necessary to acquaint students with an innovative instructional treatment. Perhaps two hours is not enough to have a variable as complex as learner control show a stable effect. Schneider (1985) indicates that students learning a complex, dynamic task show very slow acquisition rates, and that initial performance of such tasks is highly unstable. However, some aspects of subjects' behavior during practice and the immediate test may be useful to instructional developers and future experiments on learner control.

The major contributor to the treatment by test occasion interaction appeared to be the LC treatment. Subjects in this group executed the target location task more quickly during the immediate test, but performed at about the same rate during the delayed test a week later. This may have resulted from LC subjects setting the challenge level lower for accuracy than LCA subjects and LC subjects receiving more accuracy points during practice. By setting the accuracy challenge at a lower level, LC subjects seldom received negative feedback, i. e., points being subtracted from their score. The program was designed to provide more points possible when challenge levels were set higher but this also created a greater risk of losing points. For example, a subject could gain 100 accuracy points for locating a target at the highest accuracy challenge level, but the response had to be extremely accurate. If it was not, the subject risked losing 100 points. Conversely, only 10 accuracy points could be gained at the lowest level, but the response did not have to be nearly as accurate to do so. Having set the challenge for accuracy at higher levels, LCA subjects lost more points than LC subjects and probably became more cautious, and therefore, slower. LC subjects continued to respond more quickly during the immediate test, but this increased rate of speed did not continue during the delayed test.

The fact that points gained and lost appeared to change subjects' behavior during practice and the immediate test has some implications for instructional design. If students' performance is reflected in a score that is continuously updated and displayed during a practice activity, it may be possible to focus students' attention on particular aspects of the task (e. g., speed). The criterion used to keep score could be adjusted to shift each student's emphasis from speed to accuracy or vice versa as needed.

Subjects in the LCA group likely changed the levels of challenge for speed and accuracy more often than LC subjects because of the additional advice given during practice. Although all subjects received cumulative information about their performance every five trials, only LCA subjects were reminded to adjust the challenge levels every five-trial block. Both LC and LCA subjects were taught how to change level of challenge during the introduction to the practice activity and urged to adjust challenge as they practiced, but LC subject may have forgotten about this control feature. It appears that learners need to be reminded about instructional control features frequently during a lesson. If students are not prompted to use such features, they may forget the controls exist or think of them as unimportant to the task at hand. We suspect that this would be especially likely to occur when learners are engaged in a complex, fast-moving task where part of the criterion performance is to respond as quickly as possible. Even tentative conclusions about the effect of learner control on challenge rests on the assumption that learners paid attention to and were concerned with points they received for speed and accuracy. The effects could be verified by offering subjects monetary rewards for points.

It was surprising that subjects in the computer control treatment tended to favor the control of speed by the computer, while subjects in the two learner control treatments favored learner control of speed. This may be due to subjects transferring positive feelings about the overall learning experience to the CAI. It is possible that they favored the form of instructional control they were given simply because they found the task to be enjoyable or, at least, stimulating.

The results of this experiment will be used to improve future experiments on learner control of challenge and learner control over other instructional components. Descriptive information, such as number of targets subjects located before becoming tired of the task, will help shape future research that uses the same or similar tasks as a learning outcome. The difference between subjects performance across the seven-day delay shows the value of using delayed tests to determine if treatments have influenced learning or temporarily altered subjects' performance.

Researchers who plan to investigate the acquisition of complex of motor, perceptual, and cognitive skills may wish to consider conducting at least two experiments. The first can be very useful in providing important descriptive information that may lead to an increase in power of succeeding experimental designs. It may also be helpful to employ multiple practice sessions because complex, dynamic skills can require more time to mature than learning outcomes that are frequently encountered in academic settings.

References

Avner, A., Moore, C., & Smith, S. (1980). Active external control: A basis for superiority of CBI. Journal of Computer-Based Instruction, 6(4), 115-118.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and Instructional Methods. New York: Irvington.

Goetzfried, L., & Hannafin, M. J. (1985). The effect of locus of CAI control strategies on the learning of mathematics rules. American Educational Research Journal, 22(2), 273-278.

Johansen, K. J., & Tennyson, R. D. (1983). Effect of adaptive advisement on perception in learner-controlled, computer-based instruction using a rule-learning task. Educational Communications & Technology Journal, 31(4), 226-236.

Kinzie, M. B., & Sullivan, H. J. (1989). Continuing motivation, learner control, and CAI. Educational Technology, Research, and Development, 37(2), 5-14.

Klein, J. D. (1988). The effects of student ability, locus of control, and type of instructional control on motivation and performance. Unpublished doctoral dissertation, Florida State University.

Lahey, G. F., & Boyer, D. L. (1990, May). The acquisition and retention characteristics of a critical piloting subskill. Paper presented at the meeting of the Fifth Annual Mid-Central Ergonomics Conference, Dayton, OH.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. Cognitive Science, 4, 333-369.

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. Journal of Computer-Based Instruction, 16(1), 1-10.

Ross, S. M., & Rakow, E. A. (1981). Learner control versus program control as adaptive strategies for selection of instructional support on math rules. Journal of Educational Psychology, 73(5), 745-753.

Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. Human Factors, 27(3), 285-300.

Tennyson, R. D. (1980). Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. Journal of Educational Psychology, 72(4), 525-532.

Table 1

Means: Task Time: (in seconds), Range Error (in centimeters), Azimuth and Heading Error (in degrees)

Group	Immediate Test				Delayed Test			
	Time	Rng.	Azi.	Hd.	Time	Rng.	Azi.	Hd.
LCA								
M	13.7	0.16	1.8	25.8	15.7	0.21	2.4	33.5
SD	2.76	0.03	0.38	9.42	2.92	0.15	2.07	18.56
CC								
M	14.1	0.15	1.7	23.9	14.4	0.20	2.5	32.1
SD	3.05	0.05	0.56	9.24	3.72	0.12	1.40	21.08
LC								
M	12.9	0.17	1.8	24.6	14.6	0.22	2.7	29.0
SD	2.26	0.05	0.52	7.29	2.20	0.17	2.45	14.56

Note. Lower values indicate better performance.

LCA = learner control with advisement

CC = computer control

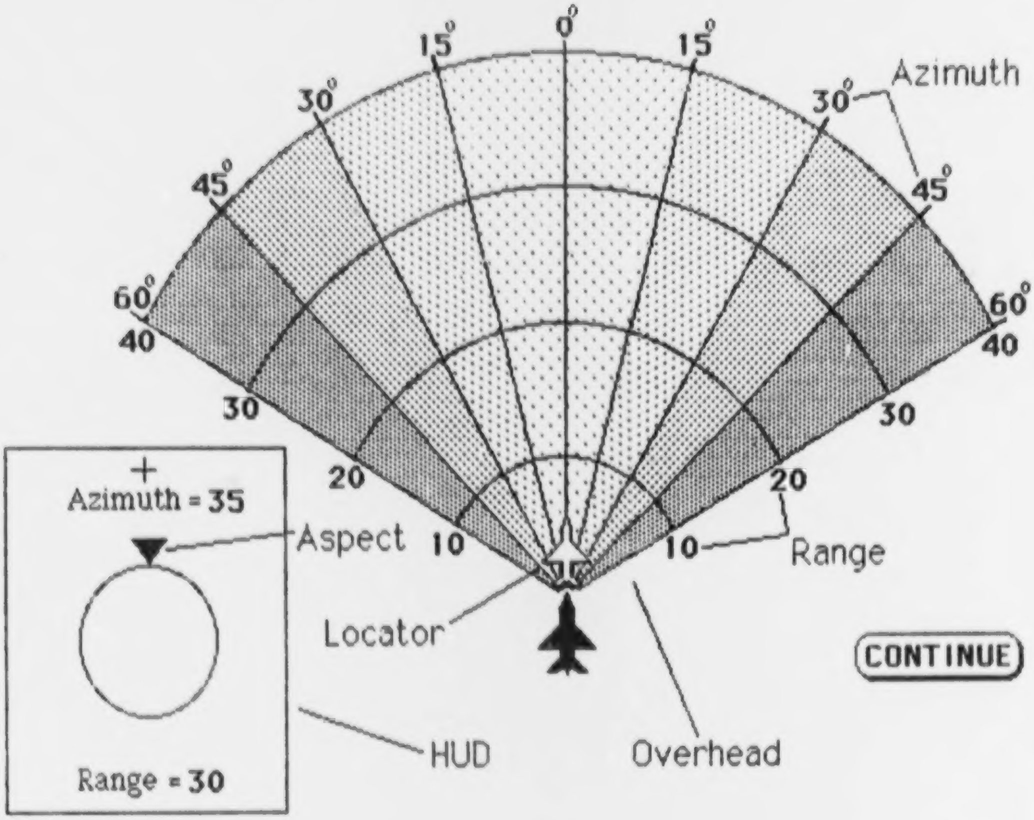
LC = learner control without advisement

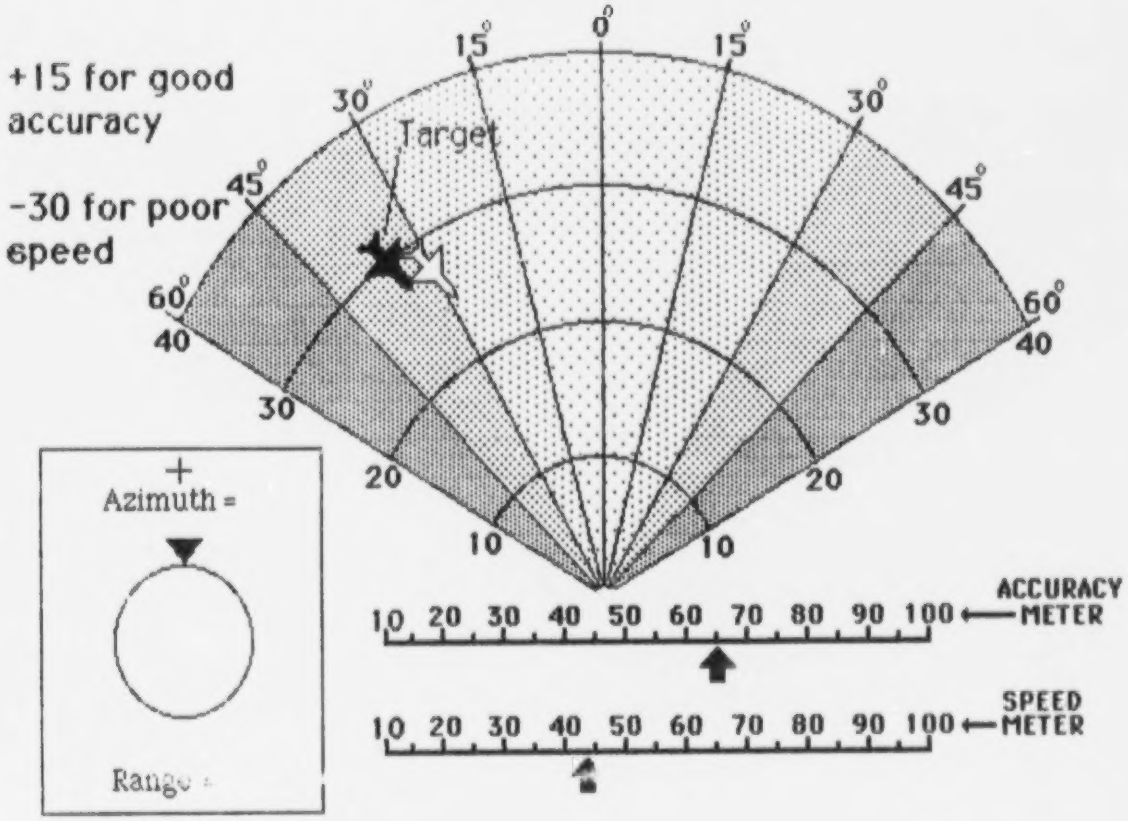
Figure Caption

Figure 1. Head-Up Display (HUD), Overhead, and target-location icons and information.

Figure 2. Points for speed and accuracy and challenge level meters.

Figure 3. Method used for answering questions and responding to bipolar attitude statements.

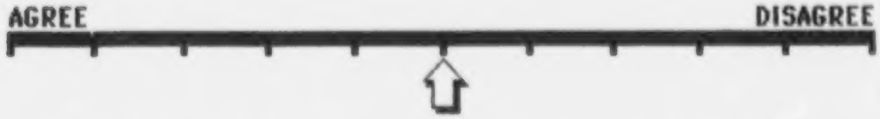

0°
15°
15°
Azimuth
30°
30°
45°
45°
60°
60°
40
40
30
30
20
20
10
10
Range
+
Azimuth = 35
Aspect
Locator
CONTINUE
HUD
Overhead
Range = 30


0°
15°
15°
30°
30°
45°
45°
60°
60°
40
40
30
30
20
20
10
10
Target
+
Azimuth =
Range =
10 20 30 40 50 60 70 80 90 100
ACCURACY METER
10 20 30 40 50 60 70 80 90 100
SPEED METER

+15 for good accuracy

-30 for poor speed

I enjoyed locating targets with the HUD Simulator.


AGREE
DISAGREE

Title:

Interactive Media Audit Trails: Approaches and Issues

Authors:

Earl R. Misanchuk
Richard Schwier

In any independent learning experience, from traditional print to computer assisted instruction, the learner strikes a path through the medium. In linear media, such as audiotaped instruction, that path is fixed and predetermined by the producer; learners share essentially similar instructional experiences. Instructional designers of such media have been able to assume a certain constancy to their instruction.

Interactive media, on the other hand, provide the learner with the opportunity to shape the program, and consequently the learning experience. Two learners may be able to traverse the same instructional materials, perhaps even achieve the same objectives, yet not share a common experience. At minimum, the degree of commonality in the experience is reduced over that experienced with the more linear media. By their very nature, interactive media are programs intentionally designed in segments, in which learner responses to structured opportunities such as menus or questions influence the sequence, size, and shape of the program (Schwier, 1987). The paths individuals strike through interactive media vary; we can assume that most learners will not share common learning experiences.

While obviously having a liberalizing (some would say humanizing) effect on instruction, the advent of interactive and hypermediated technologies brings with them a host of new questions for instructional designers, as well. Among these is the determination of the effects of taking different paths through instruction. To address such a question requires that there be some way of recording and analyzing the instructional path taken by each learner. We refer to that record of the path as the audit trail (after M. W. Petruk, personal communication, February 7, 1990).

It is a relatively simple matter to collect information about an individual's path using contemporary computer-based instruction authoring environments (e.g., *Authorware Professional* or *Course of Action*) or program development tools such as Macintosh's HyperCard. Learners' responses can be trapped and recorded at each decision point, or node. Of course, the descriptive data captured can provide very useful formative evaluation information. However, we speculate that there are additional purposes to which such information can be put, purposes that might produce generalizable rules about the construction of various paths through interactive and hypermediated instruction.

The Audit Trail

An audit trail comprises all the responses generated by a learner going through interactive or hypermediated instruction. In general terms, audit trails contain words, phrases, sentences, and paragraphs that a learner types into a computer, as well as a record of the "multiple choice"-like responses made, either from the keyboard or via some other input device, such as a mouse or touch-sensitive screen. However, to keep this preliminary discussion relatively simple, we restrict our discussion here to the "multiple choice"-like responses. Hence, we conceive of the audit trail as a string of characters (numerals or letters) representing choices made by learners as they progress through choice points—or nodes—within the instruction. For example, suppose the first point at which a learner had to make a choice had three different paths the learner could follow, and let us further suppose that she chooses the second path. As she proceeds through the subsequent instruction, she encounters a second node, with two choices, and chooses the first of those. The first two characters of her audit trail would therefore be 21 (read "two-one", not "twenty-one"). If she then chooses the first of three paths at the next node, and the third of three at the following node, her complete audit trail up to this point would be 2113. A second learner might have chosen the third path at the first node, the first path at the second node, the first path at the third node, and the second path at the fourth node, to give an audit trail of 3112.

In creating an audit trail, the computer would simply record the numerical value assigned to each option presented at a decision point, in a vector-like arrangement. As the number of decision points in the treatment increases, so will the length of the vector; as the complexity of the path taken by an individual increases (e.g., exploring optional paths as opposed to forging straight ahead), so will the length of the vector.

Audit Trails and Program Structures

Audit trails differ in nature according to the program structures in which they are generated. The structures described below represent the simplest cases; in reality, they are often combined within programs.

Linear Structure

In the most basic structure—a linear one (Figure 1a)—all learners necessarily go through the same experiences. An audit trail from a linear structure is quite simple, if not trivial. It can only contain such information as how many tries were made before the desired response was attained, and/or how long it took to attain it.

Insert Figure 1 about here.

Branching Structures

In a more complex structure—a branching one—the audit trail is correspondingly more complex, as well. In this situation, the learner is presented with choice points, or nodes, at which different responses will occasion different alternative paths through the instruction. In some structures, the paths will re-converge at the same point at which they diverged (as when a feedback loop provides remediation or supplemental information, then sends the learner back to the original node for another try) (Figure 1b); in another structure—dubbed the learner controlled parallel path, or LCPP, structure—the various (linear or branching) paths will run parallel to one another, and convergence comes somewhere down the line (Figure 1c). What distinguishes the feedback loop path from the LCPP is where the learner ends up after passing through it. In the feedback loop, the learner exits the loop back at the node of departure; in the LCPP, the learner exits the loop at some point further down the instructional sequence.

In LCPP structures, the audit trail is more complex than in the linear structure insofar as the learner's choice must be recorded. If the parallel structures are linear, then once again only the number of unsuccessful attempts and the length of time taken to criterion can be recorded in conjunction with the choice made. If, however, the branches themselves contain branches, another level of complexity obtains, and the problem of dependency shows itself: The meaning of any character in the audit trail is dependent for meaning upon the character preceding it. That is, the character '1', the second character in the first individual's audit trail, does not mean the same thing as the character '1' which is the second character in the second individual's audit trail, because the two individuals chose different paths at the first node.

In the feedback structure, a different problem is created by repeat visits to the node. An individual may or may not pass through the same node in the program more than once. Stated differently, the learner may traverse portions of a path repeatedly. If a simple frequency count is made of visits to a node, it is difficult to discriminate, for example, between the case in which two learners traversed the same path segment once each, and one in which a single learner traverses the same path segment twice. We refer to this problem as the looping problem. The looping problem is context-dependent: Interactive treatments are typically designed to permit repeated visits to some portions of instruction but not others. A second implication of this phenomenon is that not all audit trails will be of the same length.

Multimedia and Hypermedia Structures

Multimedia and hypermedia structures represent the most complex case of all, and both the dependency and looping problems manifest themselves. These structures are difficult to describe because they are so variable, both in terms of the number and kind of choices available to the learner within each display, and in terms of where the learner may go next. It is therefore not possible to draw a single generalizable flow chart for hypermedia as it is for a feedback loop or an LCPP; hence the one in Figure 1d is merely one of many possible basic structures.

We suggest that there are at least two levels of hypermediated structures: those in which cycling is permitted through relatively structured paths; and those in which paths are almost completely unstructured. While Nelson's original description of hypermedia closely matches the latter situation (Rezabeck & Ragan, 1989) that which is commonly termed multimedia matches the former description of hypermedia. The primary difference between the two levels of hypermediated structures is one of how much control is exercised by the program author and how much is given to the learner. Indeed, although the various descriptions of structures given above closely parallels their historical emergence, it also describes a continuum of control. In linear structures, the instructional designer/programmer exercises almost total control of the path of the learner (see Figure 2); in Nelsonian hypermediated structures (the second kind), almost no control is exercised by the instructional designer/programmer (except through what path or feature options are made available to the learners).

Insert Figure 2 about here.

We caution that the figure ought not to be interpreted too literally. We do not have any reason to believe, for example, that *all* LCPP structures grant half the control to the designer and half to the learner; nor can we state unequivocally that feedback loops are located exactly half-way between linear and LCPP structures. In other words, the horizontal scale should not be interpreted as being anything more than ordinal, and we are not certain that the horizontal axes representing the degree of designer control and the degree of learner control should be exactly the same length.

The audit trails of all structures described above suffer from the problem of conceptual distance. This is a classical measurement problem which surfaces in the interpretation of data from interactive treatments. Given different treatments, selections at nodes may represent nominal, ordinal or even integer data points, and each type of data imposes restrictions on how data can be analyzed. For example, one menu in a program offers the viewer a choice of "river memories," "river bridges," "river travel," or "river science." In this case "river memories" is represented as 1 and "river science" as 4. Because the data are nominal, the numerical representation is misleading; the conceptual distance implied between the two choices is in fact no greater than the conceptual distance between any two other choices from the menu. The problem is not isolated to numerical data. With tree diagrams and other graphic approaches, conceptual distance is implied as branches diverge on a diagram. In actual fact, however, choosing "A" at the seventh level of the farthest branch on the right side of a diagram is not necessarily different conceptually from choosing "A" at the fourth level of the left branch. Nevertheless, a casual observer can be seduced into thinking that spatial or numerical distance in data indicate conceptual distance as well.

Purposes of Audit Trails

To date, we have identified three distinct purposes to which audit trails can be put: as data-collection devices for formative evaluation in instructional design; as tools for basic research into the instructional design of CBI and hypermedia; and as a means of auditing usage of mediated presentations in a public forum.

Formative Evaluation in Instructional Design

The primary purpose for which audit trails have historically been used is for formative evaluation of instructional materials. It is useful to be able to determine which paths are perceived as attractive or significant by learners, and to learn where and how they make errors. If particular paths or segments of instruction receive less traffic than others, this may indicate a need for revision. In some cases, no traffic along particular paths may allow the instructional designer to eliminate those options from the system, perhaps improving the efficiency of the program or liberating space for other options.

The purpose of formative evaluation is always to optimize the performance of the product. In linear structures, learner variables are reduced in importance—they cannot be accommodated by the design. There is, of course, the option of developing parallel linear treatments for different subgroups of learners, but this is not usually a reasonable option for developers.

In branching structures, the possibility exists for either learners or designers (or both) to factor learner variables into the path decisions. The number of learner options can be increased. Efficiency is still still a goal, of course, but a greater possibility now exists for changes to be made to the instruction than in the case of linear structures.

Formative evaluation becomes less significant when used in the context of multimedia/hypermedia structures, however. Since designer influence is reduced, and learner influence is increased, efficiency ceases to exist as a construct against which to judge the performance of the materials: What does it mean to attempt to optimize the learner's path through Nelsonian hypermedia, where the learner is the only arbiter of the "correct" path? In multimedia/hypermedia, the purpose is not to optimize the treatment, but rather to open up the number of possibilities available to the learner. Formative evaluation concerns will likely be limited to cosmetic issues such as "ease of navigation" and "meaningful transitions" among elements of instruction learners encounter.

Formative evaluation is essentially a tool for instructional designers, and as structures are used which minimize the influence of instructional designers, the value of audit trails for formative evaluation purposes becomes increasingly unclear.

Basic Research in Instructional Design

Audit trails can also be used to track learner performance in research settings. Individuals, and indeed groups of individuals, can approach instruction differently, and this has traditionally been of theoretical interest. For example, consider individual differences or cognitive style constructs, such as locus of control. How might internalizers and externalizers differ in their approaches to highly organized interactive treatments? Would they react differently to, and take different paths through, very linear treatments and hypermedia treatments? One learner may select the shortest path available; another may select every available remedial segment in the same treatment. Resultant paths would be very different from one another, but, short of actually watching both individuals progress through the materials, how can these differences be expressed? As interest grows in the effectiveness of learner control of instruction (e.g., see Higginbotham-Wheat, 1990; López, 1990; Ross, Morrison, & O'Dell, 1990; Steinberg, 1977), and especially as it broadens into learner control of interactive and hypermediated instruction, these kinds of questions will command increasing interest.

Research on linear structures, as on many branching structures, is necessarily quantitative in nature. As we progress into multimedia/hypermedia structures, however, the research mode takes on a decidedly qualitative bent. On the surface, this may sound like a curious statement, but consider the following. There are relatively few questions one can ask about use of linear media beyond achievement/efficiency, performance and interactions with designs and learner variables. Meaning is imposed on the designs studied, usually by the producer or designer. For example, we could examine the effectiveness of a particular cueing strategy on different types of learners in linear media. We can ask questions like "did the cueing strategies help one group more than another?" But the treatment is fixed, so we are largely restricted to quasi-experimental designs unless parallel treatments are developed for comparison and control.

Audit trails in hypermedia fit the increasingly popular paradigm of naturalistic observation in that they are collected unobtrusively, in a natural setting. The instructional developer, given this orientation, is charged with designing a rich context within which learning can occur. Rather than being concerned about the direction and substance each learner encounters in instruction, the instructional developer is more concerned about the landscape of the instruction—the contours, breadth and depth of the terrain, and the ease with which learners can manoeuvre through the materials. The learner is viewed as part of an instructional ecosystem, simultaneously shaping and being shaped by the instruction encountered.

In their purest forms, natural interactive media are not based on instructional preconceptions, objectives or hypotheses, nor are they constructed to conform to the characteristics or needs of defined groups of learners. Rather, the learner enters the instruction unburdened, and the patterns of learning are allowed to emerge from the paths learners (and eventually groups of learners) construct. This is a potentially exciting orientation to instruction, and seems to be at odds with several of the assumptions underlying systematic instructional development (defined objectives, congruence among elements of instruction, reliable evaluation). At the same time, the instructional developer is still developing a system; this system is perhaps somewhat more organic, but its construction will still follow guidelines, conventions and rules.

One of the fundamental philosophical differences between quantitative and qualitative modes of inquiry is that a quantitative orientation emphasizes the existence of an externally definable reality. The researcher is trying to understand or reveal that reality (the absolute rules by which it operates, and how rules can be generalized). A qualitative orientation presumes the existence of multiple realities which arise from associated contexts. Given that reality is malleable and dependent upon the context in which it is observed, the researcher searches for meaning within a particular context and does not attempt to generalize meaning. Thus, meaning arises from a context, rather than by imposing meaning on a context to see if the general rules work. Since there are a huge number of potential contexts (probably finite, but who will bother counting?) multiple meanings may percolate from multiple realities.

Usage Audits for Unstructured or Public Environments

Sometimes multimedia packages are produced for use by a rather vaguely-defined audience (e.g., all visitors to a tourist site). Producers of such packages can only speculate about what content or paths in them will be most of interest and in demand, and observation of viewer/users is the only means of validating the producers' initial estimates.

A third use of audit trails therefore is to determine which paths of existing interactive media packages are of most interest to certain classes of viewers/users. The use of audit trails forms an unobtrusive way of effectively peering over the shoulders of groups of users to determine how they are traversing the interactive media package. This approach is similar to the classic unobtrusive measure of determining the amount of wear on floor tiles in front of various museum displays.

From such usage audit data, decisions might be made about optimal layouts for future treatments on videodiscs. By appropriate clustering of information, seek times may be reduced. If additional (unused) information exists on the videodisc, usage data might indicate re-vamping of the presentation is in order.

Anomalies is usage data might lead to testable research hypotheses. For example, based on observations, one might be prompted to ask the question "Do older users of a program persist in one set of activities while younger users sample from a range of activities?" As another example, in a point-of-purchase application, are people likely to make more purchases from a "motion video" display of products or from a "static video" display of products in a video catalog?

Other potential inferences include what similar subject-matter might be well received by future viewers/users (essentially needs assessment or marketing research.) If there has been heavy use of a certain class of information, it might be a clue that additional related information might also be popular.

This use of audit trails combines some of the characteristics of both formative evaluation and basic research—something of a hybrid of the earlier discussion. Certainly the goal can be to improve the effectiveness of a treatment. For instance, a point-of-purchase display might emphasize a particular type of product and de-emphasize others because of the design of selection screens. Audit trail data may suggest a need to redesign some of the selection screens. On the other hand, basic questions may arise too. For example, audit trail data may suggest that younger users consistently follow the shortest path available through certain portions of an interactive video presentation, while older users tend to linger in the same zones of the presentation. While this may ultimately have implications for formative evaluation, the data may also give rise to interesting questions of why specific design elements interact with the age of users.

In most cases, audit trail data for unstructured or public environments will include relatively simple census data. The primary interest of most users is "who is using this, and for what purpose?" The audit trail data offer information about who was exposed to which portions of a treatment, tracking user preferences rather than user performance.

Quantitative Description of the Audit Trail

In our initial search for a meaningful way to represent the audit trail, we investigated and considered several formats. The list we generated, below, is not exhaustive, but merely a point of departure. Some of the representations appear to be more useful and durable than others; some we considered briefly and discarded for various reasons outlined below.

Raw Data Matrix

The most basic way of representing audit trail data is to simply record the responses of each learner (as a vector), one above the other (see Figure 3). Matrices like these have the advantage of being easily constructed and relatively easily interpreted for each individual. The interpretation, however, can only be relatively limited, and context-bound. For crude formative evaluation purposes, the data are useful. An instructional designer can see which choices are attracting individual learners, and speculate about design decisions. But there is a serious limitation with this type of approach: raw data, by definition, aren't summarized and therefore conclusions based on the group are difficult to derive.

Insert Figure 3 about here.

Nodal Frequencies and Proportions

Another approach we investigated was to present data associated with each decision point (node) in the treatment. Data can be presented in at least two forms, as raw data (Figure 4) or as proportions (Figure 5).

Insert Figure 4 about here.

Insert Figure 5 about here.

Raw nodal frequencies, like raw data matrices, are easy to create. They are perhaps easier to interpret, since the data are now summarized. Magnitudes of differences are obvious, and at any node, comparisons have high precision and are intuitively satisfying. At the same time, relationships across nodes, or among variables are difficult to interpret. For example, how should four choices of "A" at one node be compared with 103 choices of "A" at another node, if they occupy different locations in the treatment? Perhaps only eight individuals encountered the first node, whereas several hundred encountered the second. The looping problems, described earlier in the paper, surface here to cause difficulties in interpretation. If an individual loops through a node several times, frequency data become distorted. Either the same choice is made several times, thereby inflating the frequencies at that node, or several different choices are made, thereby levelling the data at that decision point.

Nodal data can also be presented proportionally. Again, this type of format is easy to create and precision is retained at a high level. Proportions allow easier comparisons across nodes or among variables at different positions in the treatment. Of course some calculation is necessary, and the user must struggle with the question of what to use as a denominator. For example, is the denominator consistently the total number of learners encountering the treatment, or is the denominator the total number of learners who pass a particular decision point? Perhaps obviously, the denominator of any proportion will be determined by the comparisons the user chooses to make—yet another type of context dependence. As with raw nodal data, proportional data are also sensitive to looping problems. An individual looping several times through a particular node can inflate its proportion of the total. In addition, as one descends deeper into he data matrix, smaller raw numbers represent elevated proportions (see Figure 5). While proportional representation is useful for compressing large numbers, statistically bloating small numbers appears unnecessary and counterproductive.

When considering nodal data (either frequencies or proportions), individual differences are lost in the compression of data. Any design decisions based on these data are limited to conclusions about the group as a homogeneous entity, and we sacrifice any more subtle interpretations. Furthermore, data spread across several tables, each representing a single node (or, alternatively one large table showing multiple cross-breaks), are difficult to assimilate. Patterns that exist within them are difficult to detect.

Petit-Point Pattern

This early approach to the portrayal of the audit trail was suggested by some of John Tukey's work in representing non-parametric data in what he called a "stem and leaf" form (Hartwig & Dearing, 1979). It is a graphical approach, combining the intuitive appeal of a histogram with a character-based notational system symbolizing the choices. For example, in Figure 6, the X's represent the first choice, O's represent the second choice, and H's represent the third choice. A dot (•) is used to indicate that the learner proceeded past the node without making a choice, a situation made possible by an unfortunate bug in the program used to collect these sample data.

Insert Figure 6 about here.

Although virtually any symbol could be used to represent individual choices, but we chose characters that seem to occupy the same amount of space, so that inter-character differences would not influence the overall appearance, and perhaps the interpretation, of the display. The resulting pattern of characters is in some ways reminiscent of a pattern for petit-point embroidery.

To set up a display of this kind, data must be progressively or sequentially sorted decision point by decision point. That is, subsequent columns of data in a matrix must be sorted within the categories formed by the sorted data comprising the first column.

Advantages of this method included the ease of generating the display by using a search and replace function on a word processor to substitute characters, and a line sorting feature to assist in constructing the display. Although the display accurately shows the proportions of choices made at each node in a graphic and intuitive way, the dependency problem is very evidently in play: In the fifth column, for example, there are 11 distinct groups of O's; each group has a different meaning, depending on where it is located vertically. That is, although the O's all indicate that the second of the choices available was the one that was chosen, the first two O's and the third O represent choices on different content, due to the fact that different routes brought the learners to he node.

In one version of this kind of display, we also tried to use color to denote different choices, but found that it offered little advantage in interpretability.

This approach was eventually set aside because it didn't seem to do a great deal to describe what was happening. It was deemed to be moderately useful as a formative evaluation tool, but didn't appear to have sufficient power to make it useful for research purposes.

Audit Trail Tree

The audit trail tree (Figure 7) was the next approach we attempted. It combined both the graphical representation and the numerical accuracy of the Petit-Point Pattern approach, but, in addition, presented the data in a more intuitively powerful way.

Insert Figure 7 about here.

The audit trail tree is drawn so that the thickness of the line depicts the number of learners who chose the path represented. Of course, if large numbers of learners are involved, the line width could be scaled. Too, numbers (either frequencies or proportions) could be attached to each node to provide greater detail. The visual representation appeared to be useful and somewhat easier to interpret than the Petit-Point Pattern method. It was clear, it was graphical, it was intuitive, and it was grounded in reality. Comparisons were easy to make; flow could be read into the diagram as learners progressed from the beginning to the end of instruction. Although the drawing process is not difficult to do manually, it is somewhat tedious, and automation of the process on a graphic-interface microcomputer should be reasonably straightforward.

On the other hand, unless numbers were attached (as suggested above), the precision of the display was fairly low (i.e., it is difficult sometimes to tell the difference between 3 learners and 4, or between 11 and 13). The problem of conceptual distance remained, and perhaps was magnified by the ease with which other dimensions of the data were made manifest. And, of course, the problem of how to represent the loops remained.

Qualitative Description of the Audit Trail

The descriptive approaches described above are all quantitative. Descriptive analysis may also use qualitative methods productively. The goal of these descriptive approaches would be to examine the audit trail information within its context and not generalize beyond the context examined.

Content or Document Analysis

In one sense, an instructional product can be treated as a document, and a thorough examination of that document and related documentary sources can provide useful data for analysis. The usual purpose of document analysis is to explain the "status of some phenomenon at a particular time or its development over a period of time" (Best & Kahn, 1989). Any documents related to the development of an instructional product could be used in this type of analysis, and might include such things as client/developer contracts, outlines, storyboards, usage data, production schedules, other related instructional products, drafts of material, formative evaluations and the like.

One of the key cautions when using documentary sources as data is their trustworthiness. Documents , often as easily as individuals, may be inaccurate or even lack authenticity. For example, a formal contract between an instructional developer and a client might outline the parameters of a project in great detail, whereas a more accurate description of the intentions of the instructional designer might be scribbled on a cocktail napkin. Because the contract has legal and political implications, it may include cautious language which clouds the actual intentions of the contractual parties. The researcher must be careful to establish the trustworthiness of all data examined.

Some examples of audit trail studies which could employ this approach might include: to analyze the instructional design preferences of a producer of interactive media; to evaluate prejudice or bias in instructional products; or to reveal the underlying structures or levels of difficulty inherent in a particular product.

Case Study

Often used to longitudinally study development of social phenomena and change, case study methods can probably be extended to instructional products or processes. The "case" under study would not be a specific product or process, but rather an exemplar of, or prototype for, a category of products or processes. For example, a researcher conducting a case study would not be interested in a CD ROM treatment on architectural design as an entity. Rather, in a case study, the researcher might be interested in the development process and instructional design employed as an example of other products which might fall into a similar category.

Data are typically gathered from a number of sources, including direct observation, interviews, formal instruments and inventories or recorded data. The emphasis in conducting a case study is depth of analysis, not broad generalization. Generalizations would only be drawn from a series of cases which reveal consistent observations.

A number research questions in interactive media might employ such an approach. For example, a researcher interested in how people navigate through an unstructured program might choose one such program as an exemplar, and examine it in detail. Interviews with the instructional designer might reveal some of the options, limitations and assumptions inherent in the design of this particular product. Individuals using the product could be observed to see which of the navigational options are used, and which are not. Users might then be interviewed about their decisions, perhaps revealing some design reasons why certain paths were chosen or ignored. Of course, some of the quantitative approaches to describing audit trails might be useful to guide interviews, as might other formal inventories and tests. A thorough examination of this "case" from the multiple perspectives above, could reveal a great deal about the navigational preferences of individuals. Of course, the results would only be generalizable within the bounds of the characteristics of this treatment as an exemplar.

Inferential Approach to the Audit Trail

Inferential approaches would be used when comparisons are being made between groups of learners (e.g., differing on cognitive styles) or treatments (e.g., using instructors of the same gender as the learners and using instructors of differing genders).

One inferential approach to the audit trail we considered and discarded involves the use of multiple regression. This traditional approach to path analysis did not appear to be appropriate for answering the types of questions we are addressing here. It is a method of analyzing linear relationships among sets of variables, and assumes that a *causal order* among the variables is known and that the relationships among the variables are *causally closed* (Duncan, 1966). Even though we didn't give this approach much consideration, we mention it because of a possible confusion of terms: We are attempting to analyze paths through instruction, in a way that bears no relationship to path analysis, as the term is used in a statistical sense.

Furthermore, in an inferential approach to analyzing choices made in hypermediated and interactive instruction, one cannot make the assumption of normal distribution that underlies parametric statistics; hence a focus on the non-parametrics is essential.

A productive approach to analyzing a class of problems such as that under discussion would be to collect data on choices made by a large group of people, and regard that distribution as the *usual distribution* (in fact, it would be an expected distribution that is "normal" for the particular content and treatment being investigated, but since the term normal distribution has a technical connotation, we must make a distinction). Given this expected distribution, one could then subject certain individuals to treatments of varying kinds, and compare the audit trails of those subjects to the audit trails generated by the "usual" population, on a decision by decision basis (and keeping in mind the dependency problem). That is, the comparisons could only be made for single decision nodes at a time (which could be a limitation).

A statistic such as the χ^2 one-sample test, a test of goodness of fit (Siegel, 1956), would appear to be an appropriate tool for determining the statistical significance of observed deviations from "usuality". Another likely candidate would be the Kolmogorov-Smirnov test (Marascuilo & McSweeney, 1977; Siegel, 1956). Indeed, since Siegel states that "...the Kolmogorov-Smirnov test may in all cases be more powerful than...the χ^2 test" (1956, p. 51), it would seem to be the test of choice.

Advantages and Limitations of the Audit Trail Approach

Sophisticated authoring languages and systems permit an instructional developer to collect a wide range of data very easily. For example "Authorware Professional" for the Macintosh has more than 100 resident system variables and functions which can be inserted into an instructional program. It requires little sophistication to use, and resultant data can be written to files for later analysis. This is a boon for instructional developers who have specific parameters of the instruction they wish to analyze. For example, if efficiency of instruction is of interest, it is a simple matter to track the amount of time individuals devote to various instructional components, measure achievement, and from these two variables calculate an efficiency index (achievement/time). Where the instruction is stable, and relatively predictable, this offers significant opportunities to the instructional developer and researcher alike.

Of course, collecting massive amounts of data is one thing; making sense of the data is quite another. When one is conducting traditional empirical research, questions and hypotheses determine which data are important. Similarly, focussed formative evaluations which externally define the parameters for criticism and potential revision will impose limitations on the data collected. However, when the instruction is approached assumption-free, how are data excluded as insignificant? Systematically excluding data may create an impoverished description of interaction, or perhaps worse, may introduce bias into the system under investigation. It appears that the glut of data is necessary, but how can it be tamed?

In addition, unbridled data collection can result in a theoretical wasteland for individuals who are curious or eclectic (read snoopy, sloppy and unable to discard anything). Data snooping can be raised to absurd heights, and we have visions of researchers and developers paralyzed beneath mounds of imponderable data.

"The richness of nonlinear representation carries a risk of potential intellectual indigestion, a loss of goal directedness and cognitive entropy. The availability of multiple types of representations in a hypermedia system presents a cognitive overload about which little is known." (Dede, 1990, p. 20)

It is important to realize that excessive attention to detail can be dysfunctional. Often the most useful data are those which provide generalizations and trends. Extremely fine-grained information may seem like a good idea, but the most useful observations may yet be drawn from seemingly crude, but well-derived, data. While trying several of the analytical tools described in this paper, we were impressed with how quickly data grew beyond our capability to represent it. Clearly, the more complex instructional designs and questions about data become, the more we are driven into underdeveloped strategies for analysis.

Analyzing complex audit trail data meaningfully involves making inferences about two things: whether a chosen path has integrity; and why a particular path was chosen from among alternatives. First of all, we look for integrity. When an individual or group of individuals follow a particular path, we must decide whether the series of choices comprising the audit trail has any external meaning. It is possible that a series of choices may be relatively random. Therefore, in an attempt to ascribe meaning to emerging audit trails, it is possible to draw spurious conclusions. Independent observations by researchers or developers should provide a measure of reliability to conclusions.

Determining why a particular path was chosen from among alternatives is hazardous work, yet at the heart of conducting formative evaluation or basic research. Certainly post-hoc interviews and independent observation by experts can provide a measure of enlightenment; simply asking, "why did you make these choices?" can confirm suspicions. Nevertheless, individuals may choose similar paths for quite different reasons, suggesting a lack of coherence to paths, when two or more coherent cognitive paths may actually occupy the same physical geography. Of course, there is the danger that individuals confronted with the question will work to impose meaning on their own set of responses; self analysis is as natural as it is often flawed. Another potential pitfall when interviewing individuals, particularly learners, is that they have little or no understanding of the paths *not* taken. A choice of direction allows an audit trail to emerge for one group, but does not permit extensive comparison with competing audit trails for the same group. For example, an individual might say "I started here because the instruction included my name in the question, and then always chose the first option in a list after that" but can not comment on various paths which streamed away from the initial options using personal pronouns or impersonal references. When searching for meaning in audit trails, we are engaged in conceptual exploration, and it appears to us that this requires collaboration (almost conspiracy) among designers, SMEs and learners to be successful.

Issues and Challenges

One of the most pressing needs we see in audit trails research is the development of procedures for collecting and analyzing data. It is quite possible that existing procedures can be adapted to inform this area of study, but as new knowledge structures emerge, we predict that analytical tools will continue to fail to keep pace with instructional designs. Our ability to make sense of increasingly sophisticated webs of instruction depends on the development of extremely robust analytical tools and strategies. Other areas of study, such as geography or oceanography, may provide insight into ways to chart and analyze complex (and yet seemingly graphic) forms of data.

At the same time, sophisticated analytical approaches will open the door to many questions heretofore untouched. Recent work in knowledge-based management systems (KBMS) suggests research and design inquiry into multiple knowledge representation schemes and applications, improving and studying the effects of context-free information and context-sensitive advice, and the effects of immersing learners in "virtual reality" information environments (after Dede, 1990). Research in information saturated environments will allow researchers to revisit questions of learning styles and learner motivation from a fresh perspective. In a very real sense, qualitative studies will permit new questions to arise from the experiences of users within increasingly rich instructional environments. We can not afford to be arrogant at this stage of development, and suggest the range of research questions which may be important in multi-media environments. Our ability to enrich and structure information, and our continued progress in creating natural interfaces between knowledge structures and learners, far outstrip our current ability to analyze data, or even ask intelligent questions of data. We are facing an opportunity for real and meaningful exploration, and we should approach the task unfettered by some of our most fondly held assumptions (e.g.,. we can systematically manipulate instruction to adequately address the needs of learners). It is a time for listening, not necessarily answering.

References

Best, J. W., & Kahn, J. W. (1989). *Research in education* (6th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Dede, C. (1990). Knowledge processing: Navigating the sea of information. *Authorware, 3*(1), 17-22.

Duncan, O.D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72, 1-16.

Hartwig, F., & Dearing, B. E. (1979). *Exploratory data analysis.* Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-016. Beverly Hills, CA: Sage Publications.

Higginbotham-Wheat, N. (1990, February). *Learner control: When does it work?* Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Anaheim, CA.

López, C. L. (1990, February). *Personalizing math word problems.* Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Anaheim, CA.

Marascuilo, L. A., & McSweeney, M. (1977). *Nonparametric and distribution-free methods for the social sciences.* Monterey, CA: Brooks/Cole Publishing.

Rezabeck, R. H., & Ragan, T. J. (1989, February). *Elaborated resources: An instructional design strategy for hypermedia.* Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Dallas, TX.

Ross, S., Morrison, G., & O'Dell, J. (1990, February). *Uses and effects of learner control of context and instructional support in computer-based instruction.* Paper presented at the Annual Meeting of the Association of Educational Communications and Technology, Anaheim, CA.

Ryan, T. A. (1972). Multiple comparisons in psychological research. In R. E. Kirk (Ed.), *Statistical issues: A reader for the behavioral sciences* (pp. 291–306). Monterey, CA: Brooks/Cole Publishing.

Schwier, R. A. (1987). *Interactive video.* Englewood Cliffs, NJ: Educational Technology Publications.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill.

Steinberg, E. R. (1977). Review of student control in computer-assisted instruction. *Journal of Computer-Based Instruction, 3*(3), 84–90.

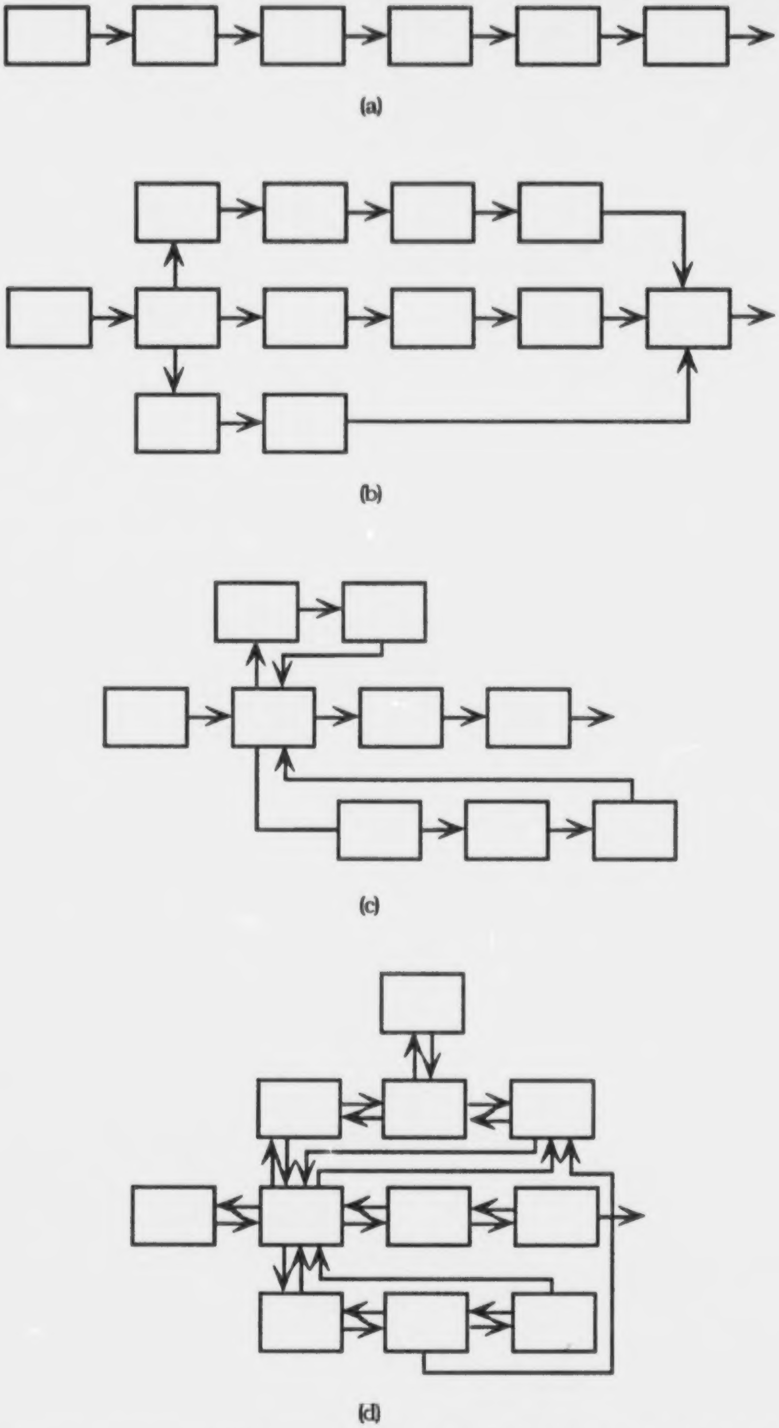


Figure 1. Program structures: (a) linear; (b) learner controlled parallel path branching; (c) classic feedback loop branching; (d) multimedia/hypermedia.

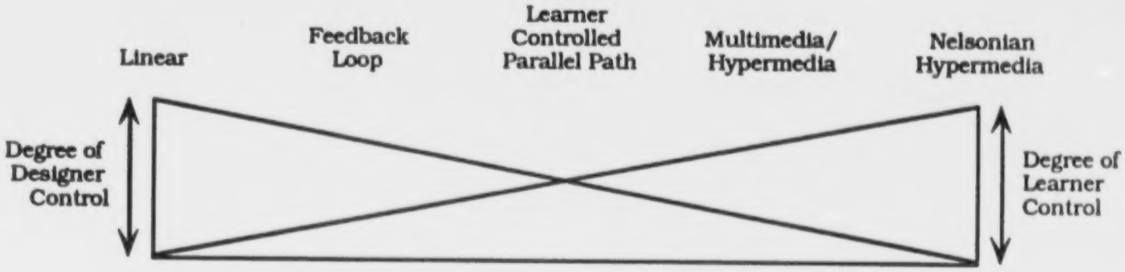

Linear
Feedback Loop
Learner Controlled Parallel Path
Multimedia/ Hypermedia
Nelsonian Hypermedia
Degree of Designer Control
Degree of Learner Control

Figure 2. Degree of designer and learner control in various structures.

Subject	1	2	2	2	2	1	1	1	1
Subject	2	2	2	2	2	1	1	1	1
Subject	3	2	2	2	2	1	1	2	1
Subject	4	2	2	2	2	1	1	2	1
Subject	5	2	1	2	2	1	1	1	1
Subject	6	1	3	2	2	2	1	1	1
Subject	7	2	2	2	2	2	1	1	1
Subject	8	2	1	2	2	2	1	1	1
Subject	9	2	3	2	2	1	1	2	2
Subject	10	2	1	2	2	1	1	2	2
Subject	11	2	3	2	2	1	1		
Subject	12	2	3	4	4	1	1		
Subject	13	2	3	4	4	1	1		
Subject	14	2	3	4	4	1	1		
Subject	15	2	2	2	2	1	1	1	
Subject	16	2	2	4	2	1	1		
Subject	17	1	1	2	2	1	1		
Subject	18	2	1	2	2	1	1		
Subject	19	2	1	2	2	1	1		
Subject	20	2	1	2	4	1	1		
Subject	21	1	3	4	1	2	1	2	
Subject	22	1	3	4	1	2	1	2	
Subject	23	1	3	4	4	2	1		
Subject	24	2	3	4	2	2	1	2	
Subject	25	2	2	2	2	2	1	1	
Subject	26	2	2	4	2	2	1	2	
Subject	27	2	1	4	2	2	1	2	
Subject	28	1	3	2	2	1	2	1	1
Subject	29	2	3	2	2	1	2	1	1
Subject	30	2	2	2	2	1	2	1	1
Subject	31	2	2	2	2	1	2	1	1
Subject	32	2	2	2	2	1	2	1	1
Subject	33	2	1	2	2	1	2	1	1
Subject	34	2	3	2	2	2	2	1	1
Subject	35	2	3	2	2	2	2	1	1
Subject	36	2	3	2	2	2	2	1	1
Subject	37	2	2	2	2	2	2	1	1
Subject	38	2	2	2	2	2	2	1	1
Subject	39	2	2	2	2	2	2	2	1
Subject	40	2	1	2	2	2	2	1	1
Subject	41	2	1	2	2	2	2	2	1
Subject	42	1	3	2	2	1	2	2	2
Subject	43	2	3	2	2	1	2	1	2
Subject	44	2	3	2	2	1	2	1	2
Subject	45	2	3	2	2	1	2	2	2
Subject	46	2	3	2	2	1	2	2	2
Subject	47	2	1	2	2	1	2	2	2
Subject	48	2	3	2	2	2	2	1	2
Subject	49	2	3	2	2	2	2	2	2
Subject	50	2	3	2	2	2	2	2	2
Subject	51	2	2	2	2	2	2	1	2
Subject	52	2	2	2	2	2	2	1	2
Subject	53	2	2	2	2	2	2	1	2
Subject	54	2	2	2	2	2	2	2	2
Subject	55	1	1	2	2	2	2	2	2
Subject	56	2	1	2	2	2	2	1	2
Subject	57	2	1	2	2	2	2	2	2
Subject	58	2	3	2	2	1	2		
Subject	59	2	3	4	2	1	2		
Subject	60	2	3	4	4	1	2		
Subject	61	2	3	4	4	1	2		
Subject	62	1	2	4	2	1	2		
Subject	63	2	2	2	2	1	2		
Subject	64	2	2	2	2	1	2		
Subject	65	2	1	2	2	1	2		
Subject	66	2	1	2	2	1	2		
Subject	67	2	1	2	2	1	2		
Subject	68	1	3	4	1	2	2	2	
Subject	69	1	3	4	2	2	2	2	
Subject	70	1	3	4	4	2	2		
Subject	71	1	3	4	4	2	2		
Subject	72	1	3	4	4	2	2		
Subject	73	2	3	2	2	2	2		
Subject	74	2	3	2	2	2	2		
Subject	75	2	3	2	2	2	2		
Subject	76	2	3	4	4	2	2		
Subject	77	2	3	4	4	2	2		
Subject	78	2	3	4	4	2	2		
Subject	79	2	3	4	4	2	2		
Subject	80	2	3	4	4	2	2		
Subject	81	2	3	4	4	2	2		
Subject	82	2	3	4	4	2	2		
Subject	83	2	3	4	4	2	2		
Subject	84	2	3	4	4	2	2		
Subject	85	2	3	4	4	2	2		
Subject	86	2	3	4	4	2	2		
Subject	87	2	3	4	4	2	2		
Subject	88	2	3	4	4	2	2		
Subject	89	2	3	4	4	2	2		
Subject	90	2	2	4	1	2	2	2	
Subject	91	1	1	2	2	2	2		
Subject	92	2	1	2	2	2	2		
Subject	93	2	1	2	2	2	2		
Subject	94	2	1	2	2	2	2		
Subject	95	2	3	2	2	1	3	2	1
Subject	96	2	1	2	2	1	3	1	1
Subject	97	2	1	2	2	1	3	2	1
Subject	98	2	1	2	2	2	3	1	2
Subject	99	2	1	2	2	2	3	2	2
Subject	100	2	2	2	2	2	4		
Subject	101	1	3	4	4	1			
Subject	102	1	2	4	1	1			
Subject	103	1	2	4	1	1			
Subject	104	1	2	4	2	1			
Subject	105	2	3	4	4	2			
Subject	106	1	2	4	2	2			
Subject	107	1	4	4	4	4			
Subject	108	1	4	4	4	4			
Subject	109	1	4	4	4	4			
Subject	110	1	4	4	4	4			
Subject	111	1	4	4	4	4			
Subject	112	2	4	4	4	4			
Subject	113	2	4	4	4	4			
Subject	114	2	4	4	4	4			
Subject	115	2	4	4	4	4			
Subject	116	2	4	4	4	4			
Subject	117	2	3	4	4	4			
Subject	118	1	2	2	2	4			
Subject	119	1	2	4	1	4			
Subject	120	1	2	4	2	4			
Subject	121	2	2	4	2	4			
Subject	122	2	2	4	2	4			
Subject	123	1	1	2	2	4			
Subject	124	1	1	2	2	4			
Subject	125	1	1	2	2	4			
Subject	126	1	1	2	4	4			
Subject	127	2	1	2	2	4			
Subject	128	2	1	2	2	4			
Subject	129	2	1	2	4	4			
Subject	130	2	1	2	4	4			
Subject	131	2	1	2	4				

Figure 3. Sample raw data matrix.

Node 1	Resp.	1																							2																							
	Freq.	32																							99																							
Node 2	Resp.	1						2						3						0						1						2						3						0					
	Freq.	7						8						12						5						28						25						41						5					
Node 3	Resp.	2			0			2			0			2			0			2			0			2			0			2			0			2			0			2			0		
	Freq.	7			0			1			7			3			9			0			5			27			1			20			5			18			23			0			5		
Node 4	Resp.	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
	Freq.	0	6	1	0	0	0	0	1	0	3	4	0	0	3	0	3	1	5	0	0	0	0	0	5	0	23	4	0	1	0	0	20	0	1	4	0	0	18	0	0	2	21	0	0	0	0	0	5
Node 5	Resp.	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ
	Freq.	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ	ǀǀǀ
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Figure 4 Tabular representation of raw frequencies of paths chosen in sample data. Although the sample data actually have eight decision nodes, only four are presented here, for simplicity. Note how quickly the size and complexity of the table grows as the number of decision nodes and/or the number of choices available at each node increases.

Node 1	Resp.	1																								2																							
	Prop.	24																								76																							
Node 2	Resp.	1						2						3						0						1						2						3						0					
	Prop.	22						25						38						15						28						25						42						5					
Node 3	Resp.	2			0			2			0			2			0			2			0			2			0			2			0			2			0			2			0		
	Prop.	100			0			13			67			25			75			0			100			96			4			80			20			44			56			0			100		
Node 4	Resp.	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2	0
	Prop.	0	86	14	0	0	0	0	100	0	42	58	0	0	100	0	33	11	56	0	0	0	0	0	100	0	85	15	0	100	0	0	100	0	20	80	0	0	100	0	0	9	91	0	0	0	0	0	100
Node 5	Resp.																																																
	Prop.																																																

Figure ~~2~~5. Tabular representation of proportions of paths chosen in sample data for first four decision nodes. Note how, as one progresses through more decision nodes, smaller frequencies of responses produce deceptively larger proportions.

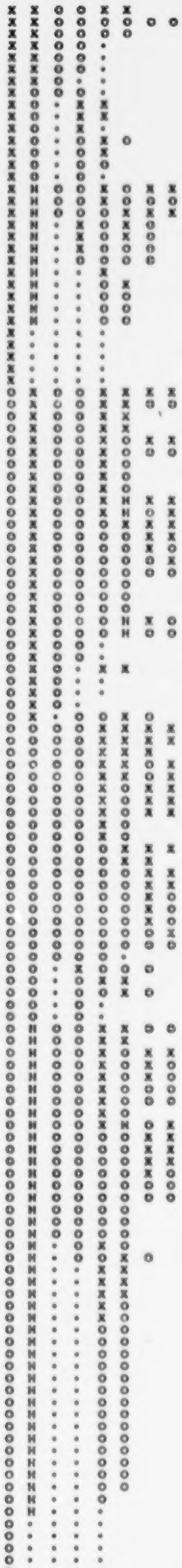


Figure 6. Petit-Point Pattern method of representing an audit trail. An X represents the choice of the first possible path from a node, an O represents the second, and an H represents the third. A • indicates no response.

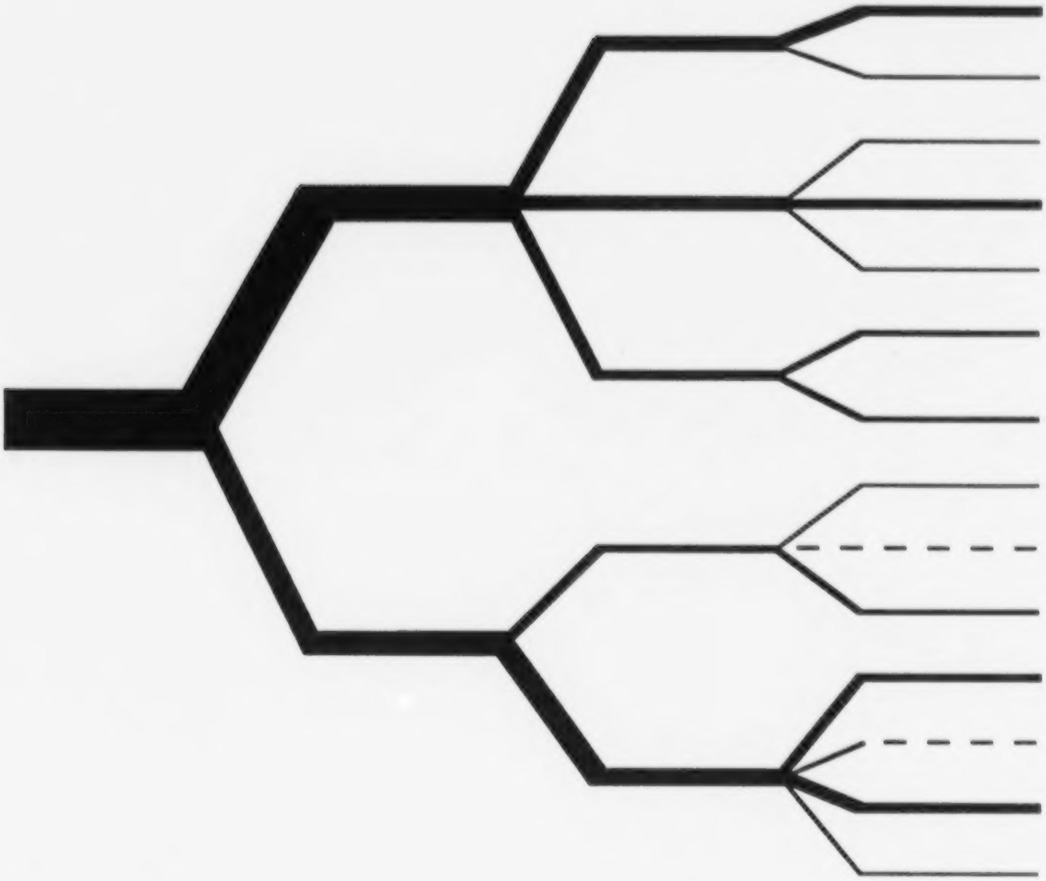


Figure 7. Audit Trail Tree of example data. The width of the line represents the number of learners taking any given path. A dashed line indicates no learners took the path.

Title:

Elementary Teacher Educational Media Utilization and Production in Selected Silicon Valley Schools

Authors:

John E. Morlan
Mei-Yan Lu

ELEMENTARY TEACHER EDUCATIONAL MEDIA UTILIZATION AND PRODUCTION IN SELECTED SILICON VALLEY SCHOOLS

John E. Morlan and Mei-Yan Lu
San Jose State University

c 1991

Introduction. Educators from the Republic of China, Taiwan; Chulalongkorn University, Thailand; and two universities in California decided it was time to up-date our knowledge concerning the utilization and production of media by teachers in selected elementary schools, grades K-6. The report that follows is based upon a study done in Silicon Valley by John E. Morlan and Mei-Yan Lu, San Jose State University, California. At a later date, comparisons will be made of media utilization by K-6 teachers in California, the Republic of China, and Thailand.

Reasons for conducting the survey. The authors of this report are professors at a large university which requires that all pre-service teachers take a three-semester unit course in instructional media and technology before a preliminary teaching certificate may be issued to the teacher education candidate. We are concerned that we not only prepare prospective teachers for the future, we must also equip them to work effectively in the classrooms of today.

The survey Instrument. No survey instrument was available which covered utilization and production of media in sufficient variety and depth. Therefore, an instrument was constructed which was considered adequate for the California study population, and could also be used in the study of media utilization and production by Thai and Chinese elementary school teachers as well. The Instructional Technology faculty at San Jose State University and the faculty of the university in Thailand, as well as selected classroom teachers, reviewed the instrument and gave suggestions for improvement prior to the printing and distribution of the questionnaire.

The Silicon Valley (California) survey. One hundred ten (110) teachers of grades K-6, representing over fifteen (15) different elementary schools, returned usable survey instruments in time for use in this summary report. In addition, eighteen (18) teachers returned forms too late to be included in the study. The survey was implemented on November 15, 1990, and returns were accepted until January 18, 1991.

Schools included in the survey. Schools represented in the survey included several schools located in low income communities, as well as schools located in middle and upper-middle class communities where "High-Tech" research, development and manufacturing companies of Silicon Valley are located.

Statistical treatment. For the purposes of this report, it was determined that a ranking of items in terms of teacher utilization of media at each grade level would be appropriate, using average weighted value for each item in the questionnaire. Items were then categorized and placed into related groupings. Frequency of media use or production activity is presented within each of these groupings. Data are presented for use of media by K-3 teachers; grade 4-6 teachers; and all teachers included in the survey, grades K-6. In addition, the Spearman-Brown rank-order correlation formula was used to determine correlation of rankings of media use by primary (K-3) and intermediate (4-6) teachers.

A grade-by-grade analysis has been completed by the researchers; however, the results of this analysis are not included in this summary report.

Figure 1: RANK ORDER OF TEACHER USE OF MEDIA, GRADES K-3

Key to Numerical Values: Daily Use - 5.0; Weekly Use - 4.0; Monthly use - 3.0
Yearly Use - 2.0; Not used - 1.0

Used daily to weekly: Bold Outline
Used weekly to monthly: **Bold**
"Top 50%" Item: *

Microcomputers

Microcomputers for reading/language arts/English as a Second Language	**3.59***
Microcomputers for math/science/health	**3.28***
Desk-top publishing, graphics (newsletters, notices, hand-outs, etc.)	2.17
Microcomputers for social studies	1.68
Microcomputers for other areas (art, music, etc.)	1.59
Interactive computer/laser disk, multimedia CD ROM	1.21
Telecommunications (modem for communication w/others, data bases, etc.)	1.05

Television (Video) and Film

Commercially produced videos (using a video player in the classroom)	2.71*
16 mm film	2.63*
Teacher copied videos (played back on a video player in the classroom)	2.38
Broadcast TV (live) from commercial or educational stations	1.93
Teacher/student made videos (using a video camera)	1.70
Laser video disks	1.37
Closed-circuit television (produced in the school district)	1.33

Still Projected Media

Overhead transparencies, teacher/student made	**3.02***
Filmstrips with cassette sound	2.78*
Overhead transparencies, commercially produced	2.35
Silent Filmstrips	2.09
Opaque projector, for preparing charts and other visuals	1.64
Silent 2 X 2 slides, teacher/student made	1.46
Opaque projector, for projection of the actual materials	1.43
Silent 2 X 2 slides, commercially produced	1.34
Slides with cassette sound, teacher/student made	1.26
Slides with cassette sound, commercially produced	1.11

Figure 1. Media Use K-3 (Cont.)

Audio Media	
Audio cassette, commercially produced	3.63*
Record player	3.45*
Cassette tape, teacher/student made	2.61
Public address system	2.03
CD player	1.16

Models and Mock-ups, Dioramas, Maps and Globes, Puppets, Field Trips	
Maps and Globes, commercially produced	3.42*
Models and Mock-ups, teacher/student made	2.63*
Field Trips, including museums, businesses, nature walks, others	2.46
Puppets, made by the teacher/student	2.43
Puppets, purchased for story telling/reading/dramatization	2.33
Maps and Globes, teacher/student made	2.32
Models and Mock-ups, commercially produced	2.30
Dioramas (three dimensional displays), teacher/student made	2.22
Dioramas (three dimensional displays), purchased	1.42

Math Manipulatives, Games and Simulations	
Math Manipulatives, commercially produced	[illegible]
Math Manipulatives, teacher/student produced	[illegible]
Games and Simulations, commercially produced	3.54*
Games and Simulations, teacher/student made	3.53*

Charts and Posters	
Charts and Posters, teacher/student made for teaching reading/lang. arts	4.23*
Charts and Posters, purchased for teaching reading/language arts	4.15*
Charts and Posters, purchased for teaching math/science/health	3.35*
Charts and Posters, teacher/student made for teaching math/science/health	3.21*
Charts and Posters, teacher/student made for teaching social studies	2.94*
Charts and Posters, purchases for teaching social studies	2.85*
Charts and Posters, teacher/student made for teaching in other areas	2.77*
Charts and Posters, purchased for teaching in other areas (music, art, sports, etc.)	2.43

Figure 1. Media Use K-3 (Cont.)

Display Boards

Blackboards (chalkboard) with chalk	4.73*
Bulletin boards/teaching displays, teacher/student made	4.11*
Bulletin boards/teaching displays, purchased	3.22*
White boards with felt markers	2.79*
Chalkboard tools: teacher made	2.60
Felt or flannel boards and materials, commercially made	2.34
Felt or flannel boards and materials, teacher/student made	2.33
Magnetic boards and magnetic board materials, commercially produced	2.02
Magnetic boards and magnetic board materials, teacher/student made	1.48

Graphic and Photographic Media Production, and Paper Materials Duplication

Hand-lettering for graphics, displays (using felt pens, chalk, etc.)	4.37*
Paper Materials Duplication: Thermal-type (Xerox) copier	4.31*
Paper Materials Duplication: Ditto-type (spirit) duplicator	4.11*
Laminating with heat lamination plastic film	3.03*
Pre-cut paper letters for graphics, displays (commercially produced)	2.62*
Rubber cement mounting (on cardboard)	2.14
Lettering tools and devices for graphic materials	2.08
Photography for producing prints	2.01
Photography for producing slides	1.61
Dry mount press used for mounting pictures	1.31
Silk screen printing on cloth, paper or cardboard	1.24
Wet mounting on cloth	1.02

Centers

Library/Book Center: Individual classroom	3.80*
Library/Book Center: For the entire school	3.69*
Interest/Discovery Center: Individual classroom	3.52*
Self-Instructional Learning Center: Individual classroom	3.42*
Self-Instructional Learning Center: For the entire school	2.01

Learner Arrangements. Learner arrangements were treated separately, with utilization frequency and rankings done outside of the over-all data treatment.

Media used in "large-group" presentations	3.54*
Media used in "small-group" instruction	2.94*
Media used in "cooperative learning"	2.47
Media used in "individual self-study" situations	2.41

Figure 2: RANK ORDER OF TEACHER TEACHER USE OF MEDIA, GRADES 4-6

Key to Numerical Values: Daily Use - 5.0; Weekly Use - 4.0; Monthly use - 3.0 Yearly Use - 2.0; Not used - 1.0

Used daily to weekly:	Bold Outline
Used weekly to monthly:	**Bold**
"Top 50%" Item:	*

Microcomputers

Microcomputers for reading/language arts/English as a Second Language	**3.03***
Microcomputers for math/science/health	2.95*
Microcomputers for social studies	2.86*
Desk-top publishing, graphics (newsletters, notices, hand-outs, etc.)	2.16
Microcomputers for other areas (art, music, etc.)	2.08
Telecommunications (modem for communication w/others, data bases, etc.)	1.31
Interactive computer/laser disk, multimedia, CD ROM	1.07

Television (Video) and Film

Broadcast TV (live) from commercial or educational stations	2.79*
Teacher copied videos (played back on a video player in the classroom)	2.72*
Commercially produced videos (using a video player in the classroom)	2.63*
16 mm film	2.44
Teacher/student made videos (using a video camera)	1.84
Closed-circuit television (produced in the school district)	1.45
Laser video disks	1.11

Still Projected Media

Overhead transparencies, teacher/student made	**3.76***
Overhead transparencies, commercially produced	**3.20***
Filmstrips with cassette sound	2.24
Opaque projector, for preparing charts and other visuals	1.81
Silent Filmstrips	1.80
Opaque projector, for projection of the actual materials	1.68
Silent 2 X 2 slides, commercially produced	1.53
Slides with cassette sound, commercially produced	1.36
Slides with cassette sound, teacher/student made	1.29
Silent 2 X 2 slides, teacher/student made	1.28

Figure 2. Media Use 4-6 (Cont.)

Audio Media	
Audio cassette, commercially produced	3.34*
Public address system	2.52*
Record player	2.40
Cassette tape, teacher/student made	2.30
CD player	1.23

Models and Mock-ups, Dioramas, Maps and Globes, Puppets, Field Trips	
Maps and Globes, commercially produced	3.54*
Maps and Globes, teacher/student made	2.71*
Field Trips, including museums, businesses, nature walks, others	2.49
Models and Mock-ups, teacher/student made	2.08
Dioramas (three dimensional displays), teacher/student made	2.05
Puppets, made by the teacher/student	1.63
Dioramas (three dimensional displays), purchased	1.62
Models and Mock-ups, commercially produced	1.52
Puppets, purchased for story telling/reading/dramatization	1.51

Math Manipulatives, and Games and Simulations	
Math Manipulatives, commercially produced	3.42*
Games and Simulations, teacher/student made	3.07*
Games and Simulations, commercially produced	2.97*
Math Manipulatives, teacher/student produced	2.85*

Charts and Posters	
Charts and Posters, teacher/student made for teaching reading/lang. arts	3.07*
Charts and Posters, purchased for teaching math/science/health	3.15*
Charts and Posters, purchased for teaching reading/language arts	2.98*
Charts and Posters, teacher/student made for teaching math/science/health	2.73*
Charts and Posters, teacher/student made for teaching social studies	2.73*
Charts and Posters, purchased for teaching social studies	2.67*
Charts and Posters, teacher/student made for teaching in other areas	2.33
Charts and Posters, purchased for teaching in other areas (music, art, sports, etc.)	2.16

Figure 2. Media Use 4-6 (Cont.)

Display Boards

Blackboards (chalkboard) with chalk	4.01*
Bulletin boards/teaching displays, teacher/student made	3.81*
Bulletin boards/teaching displays, purchased	3.22*
White boards with felt markers	2.80*
Chalkboard tools: teacher made	2.21
Felt or flannel boards and materials, commercially made	1.47
Magnetic boards and magnetic board materials, commercially produced	1.33
Felt or flannel boards and materials, teacher/student made	1.32
Magnetic boards and magnetic board materials, teacher/student made	1.32

Graphic and Photographic Media Production, Paper Materials Duplication

Paper Materials Duplication: Ditto-type (spirit) duplicator	4.42*
Paper Materials Duplication: Thermal-type (Xerox) copier	4.33*
Hand-lettering for graphics, displays (using felt pens, chalk, etc.)	3.37*
Laminating with heat lamination plastic film	2.96*
Pre-cut paper letters for graphics, displays (commercially produced)	2.59*
Rubber cement mounting (on cardboard)	1.89
Lettering tools and devices for graphic materials	1.83
Photography for producing prints	1.56
Dry mount press used for mounting pictures	1.49
Silk screen printing on cloth, paper or cardboard	1.23
Photography for producing slides	1.20
Wet mounting on cloth	1.06

Centers

Library/Book Center: Individual classroom	3.80*
Library/Book Center: For the entire school	3.69*
Interest/Discovery Center: Individual classroom	3.52*
Self-Instructional Learning Center: Individual classroom	3.42*
Self-Instructional Learning Center: For the entire school	2.01

Learner Arrangements. Learner arrangements were treated separately, with utilization frequency and rankings done outside of the over-all data treatment.

Media used in "large-group" presentations	3.49*
Media used in "cooperative learning"	3.22*
Media used in "individual self-study" situations	2.85*

Figure 3: RANK ORDER OF TEACHER USE OF MEDIA, GRADES K-6

Key to Numerical Values: Daily Use - 5.0; Weekly Use - 4.0; Monthly use - 3.0
Yearly Use - 2.0; Not used - 1.0

Used daily to weekly: Bold Outline
Used weekly to monthly: **Bold**
"Top 50%" Item: *

Microcomputers	
Microcomputers for reading/language arts/English as a Second Language	**3.31***
Microcomputers for math/science/health	**3.12***
Microcomputers for social studies	2.27
Desk-top publishing, graphics (newsletters, notices, hand-outs, etc.)	2.17
Microcomputers for other areas (art, music, etc.)	1.84
Telecommunications (modem for communication w/others, data bases, etc.)	1.18
Interactive computer/laser disk, multimedia CD ROM	1.14
Television (Video) and Film	
Commercially produced videos (using a video player in the classroom)	2.67*
Teacher copied videos (played back on a video player in the classroom)	2.55*
16 mm film	2.44*
Broadcast TV (live) from commercial or educational stations	2.36
Teacher/student made videos (using a video camera)	1.77
Closed-circuit television (produced in the school district)	1.39
Laser video disks	1.24
Still Projected Media	
Overhead transparencies, teacher/student made	**3.39***
Overhead transparencies, commercially produced	2.78*
Filmstrips with cassette sound	2.51*
Silent Filmstrips	1.94
Opaque projector, for preparing charts and other visuals	1.72
Opaque projector, for projection of the actual materials	1.55
Silent 2 X 2 slides, commercially produced	1.44
Silent 2 X 2 slides, teacher/student made	1.37
Slides with cassette sound, teacher/student made	1.37
Slides with cassette sound, commercially produced	1.28

Figure 3. Media Use K-6 (Cont.)

Audio Media

Audio cassette, commercially produced	3.49*
Record player	2.92*
Cassette tape, teacher/student made	2.45*
Public address system	2.27
CD player	1.20

Models and Mock-ups, Dioramas, Maps and Globes, Puppets, Field Trips

Maps and Globes, commercially produced	3.48*
Maps and Globes, teacher/student made	2.51*
Field Trips, including museums, businesses, nature walks, others	2.47*
Models and Mock-ups, teacher/student made	2.36
Dioramas (three dimensional displays), teacher/student made	2.13
Puppets, made by the teacher/student	2.03
Puppets, purchased for story telling/reading/dramatization	1.92
Models and Mock-ups, commercially produced	1.91
Dioramas (three dimensional displays), purchased	1.52

Math Manipulatives, and Games and Simulations

Math Manipulatives, commercially produced	3.98*
Math Manipulatives, teacher/student produced	3.52*
Games and Simulations, teacher/student made	3.33*
Games and Simulations, commercially produced	3.25*

Charts and Posters

Charts and Posters, teacher/student made for teaching reading/lang. arts	3.64*
Charts and Posters, purchased for teaching reading/language arts	3.56*
Charts and Posters, purchased for teaching math/science/health	3.25*
Charts and Posters, teacher/student made for teaching math/science/health	2.97*
Charts and Posters, teacher/student made for teaching social studies	2.84*
Charts and Posters, purchases for teaching social studies	2.76*
Charts and Posters, teacher/student made for teaching in other areas	2.55*
Charts and Posters, purchased for teaching in other areas (music, art, sports, etc.)	2.29

Figure 3. Media Use K-6 (Cont.)

Display Boards

Blackboards (chalkboard) with chalk	4.38*
Bulletin boards/teaching displays, teacher/student made	3.96*
Bulletin boards/teaching displays, purchased	3.24*
White boards with felt markers	2.80*
Chalkboard tools: teacher made	2.40
Felt or flannel boards and materials, commercially made	1.90
Felt or flannel boards and materials, teacher/student made	1.80
Magnetic boards and magnetic board materials, commercially produced	1.68
Magnetic boards and magnetic board materials, teacher/student made	1.44

Graphic and Photographic Media Production, Paper Materials Duplication

Paper Materials Duplication: Thermal-type (Xerox) copier	4.32*
Paper Materials Duplication: Ditto-type (spirit) duplicator	4.27*
Hand-lettering for graphics, displays (using felt pens, chalk, etc.)	3.87*
Laminating with heat lamination plastic film	3.00*
Pre-cut paper letters for graphics, displays (commercially produced)	2.60*
Rubber cement mounting (on cardboard)	2.01
Lettering tools and devices for graphic materials	1.95
Photography for producing prints	1.79
Photography for producing slides	1.41
Dry mount press used for mounting pictures	1.40
Silk screen printing on cloth, paper or cardboard	1.23
Wet mounting on cloth	1.09

Centers

Library/Book Center: Individual classroom	4.00*
Library/Book Center: For the entire school	3.80*
Self-Instructional Learning Center: Individual classroom	3.10*
Interest/Discovery Center: Individual classroom	3.06*
Self-Instructional Learning Center: For the entire school	2.13

Learner Arrangements. Learner arrangements were treated separately, with utilization frequency and rankings done outside of the over-all data treatment.

Media used in "large-group" presentations	3.51
Media used in "small-group" instruction	3.12
Media used in "cooperative learning"	2.84
Media used in "individual self-study" situations	2.63

CORRELATION OF PRIMARY GRADE TEACHER MEDIA UTILIZATION WITH INTERMEDIATE GRADE TEACHER MEDIA UTILIZATION

The Spearman-Brown rank-order correlation formula was used to compare the utilization practices of primary teachers with that of the intermediate teachers. A rank-order correlation of .24 was found to exist, indicating that there were significant differences in the frequency of utilization of certain media in the two groups of teachers included in the study.

CONCLUSIONS AND RECOMMENDATIONS

Preliminary examination of the data indicates that many introductory media and technology courses, and texts written to support these courses, may need revision if the needs of the beginning elementary teacher are to be sufficiently met. We need to prepare future teachers to deal with what is, while preparing them for what is to come.

Additional studies need to be done with larger populations of teachers representing other regions of the country. Middle school teachers and secondary school teachers also need to be surveyed, to determine the use of the media in their classes. Finally, a study of the use of educational media and technology by college instructors needs to be undertaken. These studies should help us as we revise our courses to meet the needs of our pre-service and in-service teachers.

Additional analysis of data shall be done during the 1991 Spring Semester. Utilization of media by K-6 teachers in Thailand and China will be compared, and a report prepared for distribution to interested parties.

Research report prepared for the 1991 AECT International Conference, Orlando, Florida, February 12-18, by:

Dr. John E. Morlan, Division Head and Professor
Educational Leadership and Development Division
San Jose State University
San Jose, California 95192-0076
FAX 408-924-3713
Office Phone Number: 408-924-3620

and

Dr. Mei-Yan Lu, Assistant Professor
Instructional Technology Program
Educational Leadership and Development Division
San Jose State University
San Jose, California 95192-0076
FAX 408-924-3713
Office Phone Number: 408-924-3645

APPENDIX

Teacher Rankings of Media Utilization

by

Primary Grades (K-3)
Intermediate Grades (4-6)
Total (Grades K-6)

UTILIZATION OF MEDIA BY TEACHERS GRADES K-6
MEDIA USED DAILY TO WEEKLY
PRESENTED IN RANK ORDER

Blackboards (chalkboard) with chalk	4.38*
Paper Materials Duplication: Thermal-type (Xerox) copier	4.32*
Paper Materials Duplication: Ditto-type (spirit) duplicator	4.27*
Library/Book Center: Individual classroom	4.00*

UTILIZATION OF MEDIA BY TEACHERS GRADES K-6
MEDIA USED WEEKLY TO MONTHLY
PRESENTED IN RANK ORDER

Math Manipulatives, commercially produced	3.98*
Bulletin boards/teaching displays, teacher/student made	3.96*
Hand-lettering for graphics, displays (using felt pens, chalk, etc.)	3.87*
Library/Book Center: For the entire school	3.80*
Charts and Posters, teacher/student made for teaching reading/lang. arts	3.64*
Charts and Posters, purchased for teaching reading/language arts	3.56*
Math Manipulatives, teacher/student produced	3.52*
Audio cassette, commercially produced	3.49*
Maps and Globes, commercially produced	3.48*
Overhead transparencies, teacher/student made	3.39*
Games and Simulations, teacher/student made	3.33*
Microcomputers for reading/language arts/English as a Second Language	3.31*
Games and Simulations, commercially produced	3.25*
Charts and Posters, purchased for teaching math/science/health	3.25*
Bulletin boards/teaching displays, purchased	3.24*
Microcomputers for math/science/health	3.12*
Self-Instructional Learning Center: Individual classroom	3.10*
Interest/Discovery Center: Individual classroom	3.06*
Laminating with heat lamination plastic film	3.00*

UTILIZATION OF MEDIA BY TEACHERS GRADES K-6
MEDIA USED MONTHLY TO YEARLY
PRESENTED IN RANK ORDER

Charts and Posters, teacher/student made for teaching math/science/health	2.97*
Record player	2.92*
Charts and Posters, teacher/student made for teaching social studies	2.84*
White boards with felt markers	2.80*
Overhead transparencies, commercially produced	2.78*
Charts and Posters, purchases for teaching social studies	2.76*
Commercially produced videos (using a video player in the classroom)	2.67*
Pre-cut paper letters for graphics, displays (commercially produced)	2.60*
Charts and Posters, teacher/student made for teaching in other areas	2.55*
Teacher copied videos (played back on a video player in the classroom)	2.55*
Filmstrips with cassette sound	2.51*
Maps and Globes, teacher/student made	2.51*
Field Trips, including museums, businesses, nature walks, others	2.47*
Cassette tape, teacher/student made	2.45*
16 mm film	2.44*
Chalkboard tools: teacher made	2.40
Models and Mock-ups, teacher/student made	2.36
Broadcast TV (live) from commercial or educational stations	2.36
Charts and Posters, purchased for teaching in other areas (music, art, sports, etc.)	2.29
Microcomputers for social studies	2.27
Desk-top publishing, graphics (newsletters, notices, hand-outs, etc.)	2.17
Dioramas (three dimensional displays), teacher/student made	2.13
Self-Instructional Learning Center: For the entire school	2.13
Puppets, made by the teacher/student	2.03
Rubber cement mounting (on cardboard)	2.01

UTILIZATION OF MEDIA BY TEACHERS GRADES K-6
MEDIA USED LESS THAN ONCE A YEAR
PRESENTED IN RANK ORDER

Lettering tools and devices for graphic materials	1.95
Silent Filmstrips	1.94
Felt or flannel boards and materials, commercially made	1.90
Microcomputers for other areas (art, music, etc.)	1.84
Felt or flannel boards and materials, teacher/student made	1.80
Photography for producing prints	1.79
Teacher/student made videos (using a video camera)	1.77
Opaque projector, for preparing charts and other visuals	1.72
Magnetic boards and magnetic board materials, commercially produced	1.68
Puppets, made by the teacher/student	1.63
Dioramas (three dimensional displays), purchased	1.62
Opaque projector, for projection of the actual materials	1.55
Models and Mock-ups, commercially produced	1.52
Puppets, purchased for story telling/reading/dramatization	1.51
Silent 2 X 2 slides, commercially produced	1.44
Magnetic boards and magnetic board materials, teacher/student made	1.44
Photography for producing slides	1.41
Dry mount press used for mounting pictures	1.40
Closed-circuit television (produced in the school district)	1.39
Silent 2 X 2 slides, teacher/student made	1.37
Slides with cassette sound, teacher/student made	1.37
Slides with cassette sound, commercially produced	1.28
Laser video disks	1.24
Silk screen printing on cloth, paper or cardboard	1.23
Silk screen printing on cloth, paper or cardboard	1.23
CD player	1.20
Telecommunications (modem for communication w/others, data bases, etc.)	1.18
Interactive computer/laser disk, multimedia CD ROM	1.14
Wet mounting on cloth	1.09

Title:

Interactive Video and Informal Learning Environments

Author:

Kristine A. Morrissey

Introduction

Each museum visitor walks through the door with a unique set of experiences, capabilities, preconceptions and anticipations. These differences influence the way visitors move through the museum, the things they attend to, and the ways they interpret and process the information. A family attempting to share "quality time" browses through the hands–on area, a graduate student in paleontology completes a class assignment by studying fossils from the Cambrian Period, and a teacher brings his students to encourage an appreciate of their heritage. Museums respond to these heterogeneous audiences without the benefit of external controls and strategies for monitoring learning that are available to formal learning institutions (such as grades, tests, and prerequisites); it is not surprising that visitor learning can be unpredictable and different from that which the exhibit designers imagined.

To encourage visitor learning, it is not enough to exhibit a collection in an aesthetically pleasing and scientifically correct display; the visitor needs guidance ". . . in learning what to attend to, how to attend to it, how to think about what is observed and to process relevant information" (Koran, Jr., Koran, Foster, 1989, p. 76). Interpretive media provide an interface between the visitor and the objects; they introduce, describe, and surround objects with a cultural, scientific, or historical context. They focus attention on critical attributes, cue visitors to intended learning, and elicit participation. Museums commonly use printed labels, brochures, slide tape presentations, models, dioramas, and audio and video tapes as interpretive media. Museum educators and exhibit designers are now exploring the capabilities of interactive technologies including hypermedia, multi–media, and other computer–mediated presentations such as interactive video. The capabilities of these media suggest the potential for interpretation that affords greater flexibility and control over learning to a range of visitors.

The Study

Interactive videodiscs are becoming increasingly popular in museum settings because their capabilities (including interactivity, durability and vast storage potential of laser discs, and the high quality of audio and video) seem to be an appropriate match for the instructional conditions and learning objectives of museums. This paper discusses a research project conducted at the Michigan

State University Museum which used an interactive videodisc (IVD) as an introduction to a special exhibit "Birds in Trouble in Michigan." The hardware components included a videodisc player (Pioneer 2000), a computer (Macintosh Plus), a video monitor, and a Kensington TurboMouse. Software included a HyperCard program and the videodisc "Audubon Society's VideoGuide to the Birds of North America: I."

The IVD program was designed to act as a conceptual pre–organizer. Presented before the visitor enters the exhibit, pre–organizers provide a way to organize or structure the information in the exhibit. The amount of novel stimuli in a museum can overwhelm visitors making it difficult for them to focus their attention for a sufficient amount of time for information processing to occur (Koran, Jr., Longino, and Shafer, 1983). Conceptual pre–organizers can overcome this problem by orienting visitors to what they are about to see; visitors can then "better process what they see, adopt strategies appropriate to their interests and time, and ignore details and elaborations that may cause unnecessary confusion and overload" (Screven, 1986, p. 124). The HyperCard program designed for this project used computer generated text screens and segments from the videodisc to introduce vocabulary and discuss concepts and principles illustrated in the exhibit such as "extinction" and the relationship between survival and loss of habitat (see Appendix A for sample screens from the HyperCard program). Video segments also provided a context for observing the mounts by discussing identification and illustrating bird behavior in their habitats.

Research Questions

1. "Will an interactive video program used as a conceptual pre–organizer increase visitors' time within an exhibit?" Museums are by their definition, a place that houses, displays, and studies collections. However, rather than directing visitors' attention to the exhibit, the novelty of the device and the energy and excitement of sound, color, and motion could distract attention from the exhibit. If the program was an effective conceptual pre–organizer, the study hypothesized that visitor time within the exhibit area would be increased by the addition of the program. Attention has long been used as an indicator of museum learning and elapsed time within an exhibit, the most commonly used measure of attention, has been consistently correlated with

both cognitive and affective outcomes (Falk, 1983; Koran, Jr., Foster and Koran, 1989).

2. "What types of visitors will use the program?" The term "attracting power" has been used to refer to the ability of an exhibit or device to attract the viewer. In a free-choice setting, any device or exhibit must first attract and hold visitors' attention before learning can take place. In a symposium of media comparison studies at this conference two years ago, it was argued that the learner's attitude toward a medium can influence whether he or she attends to the medium and can interact with the message to either promote or deter learning (Yacci, 1989). Visitor studies suggest that there is indeed a relationship between the type of medium and its attracting power (Screven, 1975; Beer, 1987) and the attracting power is not always related to its instructional impact. To make decisions about the design of interpretive media, museum educators must know not only which medium can support the necessary instructional strategies, but also which medium will attract the targeted audience.

This study was interested in the type of visitors the IVD system would attract. Two variables were selected as likely to influence whether visitors used the program: "composition of the group" (number of adults and/or children) and "gender" of group members. Groups were investigated rather than individuals because most visitors are part of a social group and their behavior and expectations are influenced by the members and context of the group (Dierking, 1989; Kropf, 1989).

3. "How will visitors respond to the levels of options offered by the program?" The ability to give the user control over variables such as the content, sequence, or pace of presentation is intuitively attractive to the museum educator interested in meeting the needs of a heterogeneous audience. However, research does not support the assumption that user control is always advantageous but rather suggests that learners respond differently to these options. This issue is further complicated for museums as most interactive media research has been conducted in formal settings and the results may not pertain to informal learning environments.

This study did not attempt to resolve the complex issues associated with user control variables; instead, it attempted to identify whether two variables ("interest in birds" and "computer comfort") were related to the types of choices made by visitors. Two levels of menus were offered with the second level offering more information and examples. The study hypothesized that if visitors' choices were conscious selections based on interests and abilities, there would be a relationship between their interests and their interaction with the program. Visitors with low interest in the subject would be less likely to ask for more information or examples than those with high interest. Users' comfort level with the computer might also influence the interaction, with "high–comfort" users selecting more options and "low–comfort" users confining their choices to the main menu.

Methodology

The study took place over ten weekend days with five days randomly assigned to control and five to experimental conditions. Under the experimental conditions, the interactive videodisc system was set up for a two hour period outside the entrance to the exhibit hall. On control days, the system was not present. A video camera was mounted outside the exhibit area with a view of the IVD station and one wall of the exhibit (see Figure 1). All visitors entering the exhibit area during the two hour data collection periods on control and experimental days were part of the study. A total of 366 groups were included, with 181 groups visiting on the experimental days and 185 on the control days. The number of individuals in each group ranged between 1 and 8.

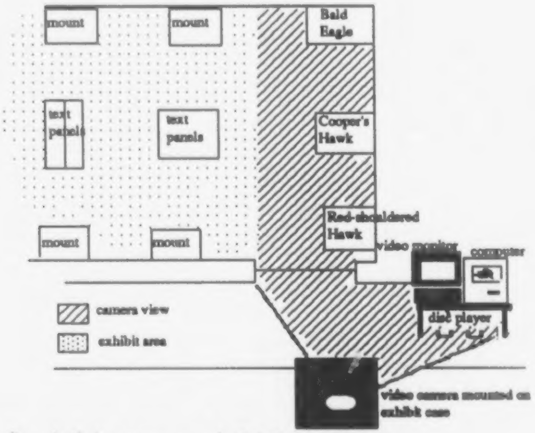

mount
mount
Bald Eagle
text panels
text panels
Cooper's Hawk
Red-shouldered Hawk
mount
mount
video monitor
computer
disc player
camera view
exhibit area
video camera mounted on exhibit case

Figure 1. Diagram of exhibit area and IVD program

The videotapes were analyzed to identify the mean time spent within the exhibit area and mean time in front of two specific bird mounts (Bald Eagle and Red-shouldered Hawk). The mean times on the control days were compared to the mean times on the experimental days (with the IVD program available outside the exhibit). All groups were coded by their composition and the gender of the adults and children. Groups who used the program were compared to all groups using chi-square. The HyperCard program recorded all selections made by each group and asked each group to identify interest in birds ("High", "Medium", "Low") and if they were comfortable with computers ("Yes", "Somewhat", "No"). Each group was coded as a Level One user if all selections were from the Main Menu and as Level Two if at least one secondary menu was used. Level One and Level Two users were compared (using chi-square) on "computer comfort" and "interest in birds."

Results

Table 1a shows the difference in mean times within the exhibit area between the control group and the experimental group (including those who used the program and those who did not.). On experimental days, the mean time of 167 seconds was significantly higher than the mean of 126 seconds on control days (ANOVA, p<.001). This suggests that the IVD program did not compete for visitors' attention but increased their time within the exhibit.

Table 1b compares the mean times in front of the Red-shouldered Hawk and the Bald Eagle mounts. On the experimental days, the average viewing time for the hawk was 3.2 seconds, which was statistically higher (ANOVA, p<.001) than the 1.2 seconds on control days. The mean time for the eagle was 8.4 seconds on the control days and 8.5 seconds on the experimental days, which was not significantly different. The differential effect cannot be readily explained; possibly visitors were less familiar with the hawk, and therefore the program had more influence on their behavior related to it.

During the experimental days, visitors who did *not* use the program spent an average of 160 seconds in the exhibit area, which was significantly more time than the 126 seconds of the control group (ANOVA, p<.004). Three possible explanations for this interesting phenomenon are: 1. Many visitors were "peripheral viewers"—watching over the shoulders of other groups but not actually using the program. They may have seen a brief segment or text screen that influenced their behavior. 2. The presence of the IVD program may have instigated associations and expectancies that resulted in increased attending behaviors; visitors may have assumed that "this exhibit is important." 3. Those who used the program became more involved with the exhibit, and their behaviors were then imitated by non–users. Modeling has been suggested as a strong factor in the behavior of museum visitors (Koran, Jr., Koran, Dierking, and Foster, 1988).

Table 1a

Elapsed Time Spent within Exhibit Area

	Groups	Mean time within exhibit
Control days	181	126 seconds*
Experimental days	185	167 seconds*
Combined	366	145 seconds

** Statistically significant, p<.001*

Table 1b

Viewing Times at Two Mounts

	Control days	Experimental days
Red–shouldered Hawk	1.2 seconds *	3.2 seconds*
Bald Eagle	8.4 seconds	8.5 seconds

** Statistically significant, p<.001*

Characteristics of Program Users

Both variables investigated ("group composition" and "gender") were significantly related to use of the program (chi–square, p<.001). Groups with children were more likely to use the program than those with adults only. Groups composed of a single adult with children showed the highest incidence of program use. This group also averaged the least time within the exhibit when the program was not present (104 seconds). Interactive programs may attract visitors who can most benefit from the guidance.

Table 2 identifies the percentages of experimental groups that used the program broken down by gender of the adults (by column) and of the children in the group (by row). Each cell identifies the percentage of the experimental groups (with that particular description) that used the program. Comparing the row for groups with girls to the row for groups with boys shows that in each of the three categories, groups with boys were at least twice as likely to use the program as groups with girls. Groups with both a male and female adult were five times more likely to use the program if they had a boy than if they had a girl (50% to 9%). The differences are not so distinct when the column for adult females is compared to that of adult males. Males were somewhat less likely to use the program than females (17% to 25%) when the child was a girl; but if the child was a boy, the gender of the adult made little difference. The highest percentage of use (83%) was observed in groups with a male adult and both a boy and a girl.

Table 2

Percentage of Groups using Grogram by Gender

	Adults		
Children	**Female(s)**	**Male(s)**	**M and F**
None	7%	13%	14%
Girl(s)	25%	17%	9%
Boy(s)	50%	55%	50%
Girl(s) and Boy(s)	20%	83%	25%

By observing families in a spontaneous and somewhat anonymous environment, this research offers a unique vantage point to study the relationship between gender and use of technology. Unfortunately, only a limited number of studies have investigated the relationship between use of technology and gender in the museum setting. In one study, more males were direct users of the computer, but both genders were equally represented as indirect users (Pawlukiewicz, Doering, and Bohling, 1989). In another study utilizing touch screen computers as an orientation device, 65% of the users were males (Sharpe, 1983). The effect of gender may vary with different audiences and content areas. In this study, the program was popular, and visitors sometimes had to wait to use it. Males may have been more likely to do so. Further research on the use of technology in informal learning environments may help educators understand families' expectations and associations related to the use of technology.

Reactions to the IVD Program

While the device in this study and the size of the monitor limited the number of users, the program seldom had only one user— more often a group was gathered around it. Families crowded onto the two available chairs with the children most often controlling the trackball and parents offering suggestions or pointing at the screen. The average time spent with this program was almost 6 minutes, with 18 minutes as the longest time; this was considerably longer than maximum of three minutes holding time that previous research has shown for other media formats (Beer, 1987). The analysis of options selected showed that 75% of the users created a unique path through the program. For example, one path included returning to the main menu twice, watching three video segments on hawks, and answering three questions related to hawks. The path of another group included using the menu once and viewing the Bald Eagle segment twice. Twenty–four groups repeated at least one video segment.

The level of users' paths was significantly related to self–assessed interest in birds (chi–square, $p<.05$). Level Two users were more likely to have a higher self–assessed interest in birds than Level One users. However, the level was not related to self–assessed computer comfort; there were no significant differences between Level One and Level Two users in the way they assessed their comfort with computers. This suggests that once visitors began to use the system, they felt comfortable interacting with the choices, regardless of their previous exposure to computers. However, as interest in birds and comfort with computers were not measured, but self–reported, these data must be interpreted cautiously.

Summary and Discussion

This study investigated visitor responses to an interactive videodisc program used as a conceptual pre–organizer for an exhibit on endangered birds. The presence of the interactive videodisc program significantly increased visitors' time within the exhibit area, suggesting increased attention and learning. Groups with children were more likely to use the program than adult only groups. Groups with males (either adult or children) were more likely to use the program than those with females. This provokes questions about the relationship between individual characteristics and the use of technology in informal learning settings.

Seventy five percent of the program users selected a unique path through the program. The level of the path (One or Two) was related to "interest in birds" but not to "comfort with computers." This suggests that visitors did take advantage of the option to create a presentation appropriate for their interests and this usage was independent of their comfort with computers.

These results do not suggest that interactive video *should* be used in a museum environment, but rather that it *can* be used effectively as a pre–organizer for an exhibit without competing with the exhibit for visitors' attention. Being able to select the content, sequence, pace, and amount of information presented, visitors can actively research their own interests, selecting from a potentially vast number of presentations. This empowering of visitors can be an important advantage for museums and other informal learning institutions which strive to encourage life–long learning for heterogeneous audiences.

References

Beer, V. (1987). Great expectations: Do museums know what visitors are doing? Curator, 30 (3), 206–215

Dierking, L. D. (1989). The family museum experience: Implications from research. Journal of Museum Education, 14 (2), 9–12

Falk, J. H. (1983). The use of time as a measure of visitor behavior and exhibit effectiveness. Roundtable Reports: The Journal of Museum Education, 7(4), 10–13

Koran, J. J., Jr., Foster, J. S., Koran, M. L. (1989). The relationship amount interest, attention and learning in a natural history museum. In S. Bitgood, A. Benefield, & D. Patterson (Eds.), Proceeding of the 1989 Visitor Studies Conference (pp. 72–79). Jacksonville, AL: Center for Social Design.

Koran, J. J., Jr., Koran, M.L., & Foster, J. S. (1989). The (potential) contributions of cognitive psychology to visitor studies. In S. Bitgood, A. Benefield, & D. Patterson (Eds.), Proceeding of the 1989 Visitor Studies Conference (pp. 72–79). Jacksonville, AL: Center for Social Design.

Koran, J. J., Jr., Koran, M. L., Dierking, L. D. & Foster, J. F. (1988). Using psychological modeling to direct attention in a natural history museum. Curator, 31 (2) 36–43

Kropf, M. B. (1989). The family museum experience: A review of the literature. Journal of Museum Educatoion, 14 (2), 5–8

Pawlukiewicz, J. D., Doering, Z. D. and Bohling, K. (1989). "The Caribou Connection: Will People Stop, Look, and Question?" Washington, DC: Institutional Studies, Smithsonian Institution

Screven, C. G. (1975). The effectiveness of guidance devices on learning. Curator, 18 (3), 219–243

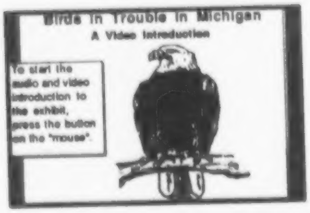
Screven, C. G. (1986). Exhibitions and information centers: Some principles and approaches. Curator, 29 (2), 109–137

Sharpe, E. (1983). "Touch–screen Computers An Experimental Orientation Device at the National Museum of American History." Washington, DC: Smithsonian Institution

Yacci, M. (1989). The singer of the song. In M. R. Simonson and D. Frey (Eds.) Proceeding of Selected Research Paper Presentations at the 1989 Annual Convention of the Association for Education Communications and Technology, pp. 140–148

APPENDIX A

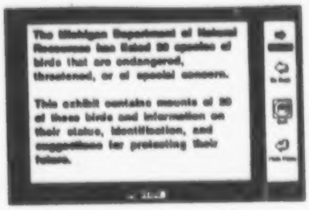
Sample Computer Screens from the Interactive Videodisc Program



Birds in Trouble in Michigan

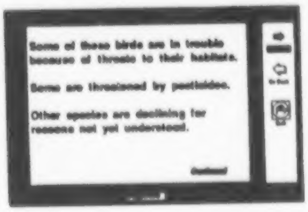
A Video Introduction

To start the audio and video introduction to the exhibit, press the button on the "mouse".



The Michigan Department of Natural Resources has listed 30 species of birds that are endangered, threatened, or of special concern.

This exhibit contains mounts of 20 of these birds and information on their status, identification, and suggestions for protecting their future.

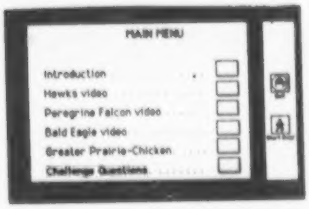


Some of these birds are in trouble because of threats to their habitats.

Some are threatened by pesticides.

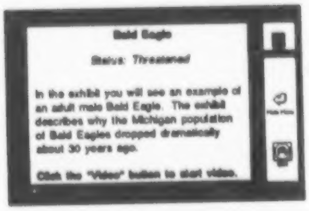
Other species are declining for reasons not yet understood.

Continued



MAIN MENU

Introduction ☐
Hawks video ☐
Peregrine Falcon video ☐
Bald Eagle video ☐
Greater Prairie-Chicken ☐
Challenge Questions ☐

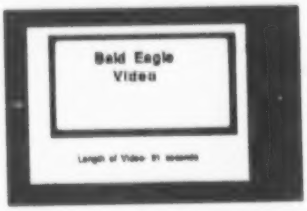


Bald Eagle

Status: *Threatened*

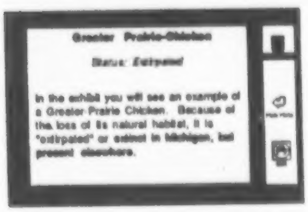
In the exhibit you will see an example of an adult male Bald Eagle. The exhibit describes why the Michigan population of Bald Eagles dropped dramatically about 30 years ago.

Click the "Video" button to start video.



Bald Eagle Video

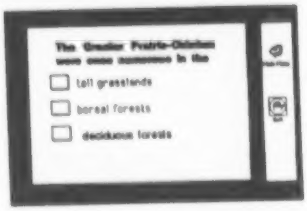
Length of Video: [illegible] seconds



Greater Prairie-Chicken

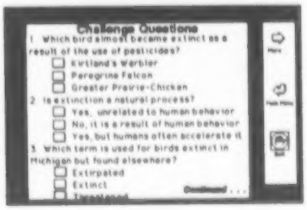
Status: *Extirpated*

In the exhibit you will see an example of a Greater Prairie Chicken. Because of the loss of its natural habitat, it is "extirpated" or extinct in Michigan, but present elsewhere.



The Greater Prairie-Chicken were once numerous in the

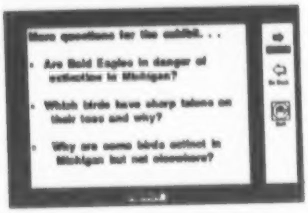
☐ tall grasslands
☐ boreal forests
☐ deciduous forests



Challenge Questions

1. Which bird almost became extinct as a result of the use of pesticides?
   - ☐ Kirtland's Warbler
   - ☐ Peregrine Falcon
   - ☐ Greater Prairie-Chicken
2. Is extinction a natural process?
   - ☐ Yes, unrelated to human behavior
   - ☐ No, it is a result of human behavior
   - ☐ Yes, but humans often accelerate it
3. Which term is used for birds extinct in Michigan but found elsewhere?
   - ☐ Extirpated
   - ☐ Extinct

Continued . . .



More questions for the exhibit. . .

- Are Bald Eagles in danger of extinction in Michigan?
- Which birds have sharp talons on their toes and why?
- Why are some birds extinct in Michigan but not elsewhere?

BEST COPY AVAILABLE

Title:

Learning Dynamic Processes from Animated Visuals in Microcomputer-based Instruction

Author:

Gary B. Mayton

The association of the use of computer-generated visual images with specific types of learning outcomes has begun to show promising results (Rieber, 1988; Baek and Layne, 1988). While researchers exploring visual form effects on a range of learning tasks have suggested graphic detail and embedded learning strategy differences (Dwyer, 1978; Belland, Taylor, Canelos, Dwyer, and Baker, 1985; and Belland, Mayton, Dwyer, and Taylor, 1991), their results consistently show diminished performance on learning tasks related to complex interactions within a system. That is, while conclusions about the effects of visual image detail and modeled strategies for viewing visual information in microcomputer-based instructional materials are drawn for a variety of performance criteria, those tasks which require the learner to recall information about dynamic processes of a presented system yield low results. The purpose for conducting the study reported here was to determine if a particular form of visual image, namely animation, (delivered via a microcomputer-based instructional system) would enhance performance on criterion measures demonstrating the learning of dynamic processes.

Animation, the rapid display of slightly varied images in order to produce a continuous sequence of an object(s) in motion, brings a special combination of potentials to the visual world. Producing a moving sequence of images enables one to add the fourth dimension of time to a visual scene. Such an attribute allows the depiction of multivariate changes in an event(s) over time. Also, individual animation images (frames) are typically created through hand-drawn, modeled, sculpted, or computer-rendered means. Such a process affords a special attention to detail and precision. For similar reasons, the animation process permits the rendering of scenes that could be difficult, dangerous, or impossible to depict by traditional photographic, cinematic, or videographic techniques. In light of this combination of attributes, it is anticipated here that animation would be instructionally attractive for presenting information about dynamic processes.

This study was conducted to investigate the effects of the animation of visual information when used in conjunction with text and static visuals in microcomputer-based instruction. Of particular interest was consequent performance on higher level learning tasks as a result of viewing these forms of visual presentation. The specific interest was on those levels of learning that demonstrate knowledge of a dynamic process.

In addition to comparisons of contrasting forms of instructional presentations (i.e., animated versus static visuals), attention was also paid to some other presumably relevant factors in this study. Focus on the learning of dynamic processes from animation-enhanced instruction also included attention to forms of microcomputer-based materials which did and did not employ an externally-imposed, visualization strategy, to testing in both graphic- and text-based formats, and to retention of learning.

Hypotheses

Stated in the null form, the following hypotheses were tested:

***Ho*$_1$.** Persons exposed to animated graphics (used with static visual images and text) in microcomputer-based instruction will not score significantly higher than persons receiving microcomputer-based instruction using static visual images and text alone on cued-recall tests of dynamic processes and related complex concepts when the performance measures are taken immediately following presentation of instruction.

Ho2. Persons exposed to animated graphics (used with static visual images and text) in microcomputer-based instruction will not score significantly higher than persons receiving microcomputer-based instruction using static visual images and text alone on cued-recall tests of dynamic processes and related complex concepts when the performance measures are taken following a one-week delay after presentation of instruction.

Ho3. Persons exposed to animated graphics (used with static visual images and text) in microcomputer-based instruction will not score significantly higher than persons receiving microcomputer-based instruction using static visual images and text alone on free-recall tests of dynamic processes and related complex concepts when the performance measures are taken immediately following presentation of instruction.

Ho4. Persons exposed to animated graphics (used with static visual images and text) in microcomputer-based instruction will not score significantly higher than persons receiving microcomputer-based instruction using static visual images and text alone on free-recall tests of dynamic processes and related complex concepts when the performance measures are taken following a one-week delay after presentation of instruction.

Design

To investigate these hypotheses, 72 introductory psychology, undergraduate students at The Ohio State University participated in a 3X2X5X2 repeated measures, factorial design study in the Autumn of 1989 and Winter of 1990. The independent variables included the three levels of instructional/presentation method employed and two visual orientations of the subsequent criterion measures. The dependent variables included the five criterion measures. In order to investigate retention of learning, these tests were administered twice, immediately after treatment exposure and following a one-week delay. Figure 1 shows the variables represented in the study. Subjects were randomly assigned to one of three treatment groups to complete a microcomputer-based tutorial presentation of information about the structure and function of the human heart. Random assignment was made again for the purpose of grouping the subjects of each treatment group for one of the two parallel forms of testing. The five performance measures were administered at that time and again following the one-week delay. Experimental arrangement is diagrammed in Figure 2.

Treatments

In each of the three treatment groups, subjects were presented information about the human heart via a combination of text and graphic images. Content and sequence for this presentation were based on materials developed by F. Dwyer (1972, 1978) for a line of studies on visual learning. Each subject viewed 54 frames (screens) of information about location and function of parts of the human heart as well as overall heart function and cycling. The text for each of the 54 frames consisted of one to four sentences (approximately ten to sixty words). A sample information screen is shown in Figure 3. The graphic images consisted of a combination of static, two-dimensional line drawings of the heart, text labels, and arrows. In each frame, the graphic depicted a lateral, posterior section of the human heart. The detail of the graphic was consistent with the form of visual generally found most effective by Dwyer [1978]. Labels consisted of a one- or two-word identifier while arrows were used to indicate a part, process, or flow.

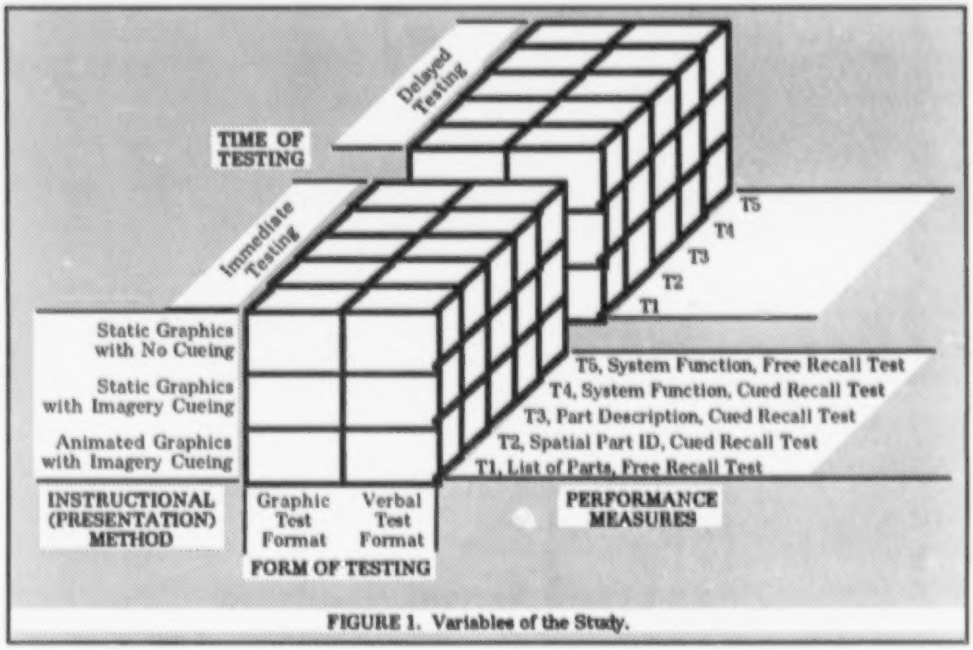

TIME OF TESTING
Delayed Testing
Immediate Testing
T5
T4
T3
T2
T1
Static Graphics with No Cueing
Static Graphics with Imagery Cueing
Animated Graphics with Imagery Cueing
T5, System Function, Free Recall Test
T4, System Function, Cued Recall Test
T3, Part Description, Cued Recall Test
T2, Spatial Part ID, Cued Recall Test
T1, List of Parts, Free Recall Test
INSTRUCTIONAL (PRESENTATION) METHOD
Graphic Test Format
Verbal Test Format
FORM OF TESTING
PERFORMANCE MEASURES

FIGURE 1. Variables of the Study.

The information frames were interspersed with 57 questions concerning the content. Computerized assessment of each response provided immediate feedback to the subject in the form of a simple message acknowledging response accuracy. When a response was incorrect, the subject received the option to repeat the specific information frames that related to that question and a second opportunity to correctly answer the question.

Treatment One: Static Visuals, No Cueing Strategy. The first treatment group received this microcomputer-based tutorial on the human heart, employing text and static visuals interspersed with computer-assessed, short-answer questions as described above.

Treatment Two: Static Visuals, Imagery Cueing Strategy. The second treatment included the same presentation features as the first with the addition of an external learning strategy developed by Taylor, Belland, Canelos, Dwyer, and Baker [1987]. This imagery cueing strategy involves the periodic inclusion (every five frames of presentation) of a special frame which shows an outline of the human heart and requests that the viewer fill in parts and operation of the heart mentally. It was anticipated that this would encourage viewers to mentally visualize the information presented, thus emphasizing the processing of visual information and enhancing the learning of the content. The second treatment, therefore, is identical to the first (static graphics and text) with the addition of the imagery cueing, learning strategy.

Treatment Three: Animated Visuals, Imagery Cueing Strategy. The third treatment preserved the content, sequence, and operation of the second with one exception. Whereas the second treatment employed static visuals, the third treatment also included animated sequences of a pumping heart, where appropriate. The animated visuals were appropriately incorporated in the 54 frames of information described above to depict changes in shape, flow of blood, and movement of parts

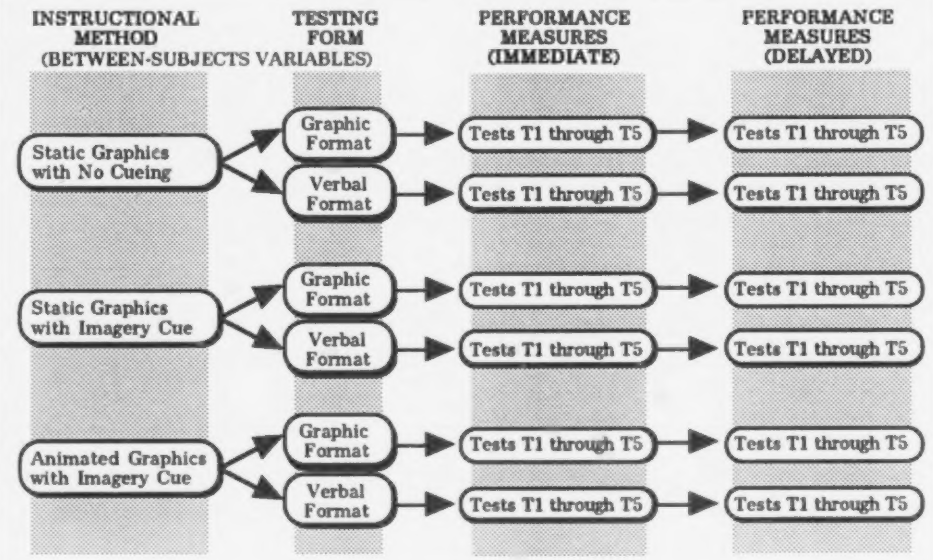

INSTRUCTIONAL METHOD
TESTING FORM
(BETWEEN-SUBJECTS VARIABLES)
PERFORMANCE MEASURES (IMMEDIATE)
PERFORMANCE MEASURES (DELAYED)
Static Graphics with No Cueing
Static Graphics with Imagery Cue
Animated Graphics with Imagery Cue
Graphic Format
Verbal Format
Tests T1 through T5

FIGURE 2. Experimental Arrangement.

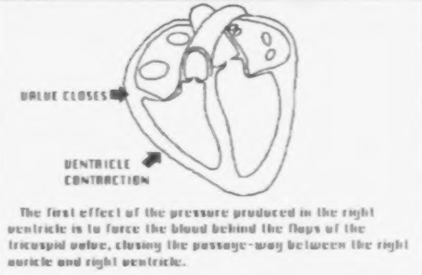
within the heart. The animation consisted of the cycling of an entire set of 16 images of the human heart or an appropriate subset of the 16 images.

Rendered detail of the graphics for both treatments was identical (i.e., two-dimensional, non-shaded, line drawings of the heart in cross-section). But, in those instances when the text described a process over time, the visual became animated to depict the information over the added dimension.

Criterion Measures

Following exposure to the treatments, each subject was randomly assigned to one of two testing groups for completion of a set of five tests measuring performance on a variety of learning tasks. One testing group completed the five tests written in a format dominant in a verbal orientation, while the other group completed a parallel set of tests which emphasized stronger reliance on the graphic elements This graphic form of test was developed and validated by C. Dwyer (1984). The five tests measured (a) list learning tasks of cardiac part names (in a free recall format); (b) spatial, cued-recall tasks of part identifications; (c) simple concept learning tasks (description of parts, cued-recall); (d) complex concept learning tasks (cued recall of system operations); and (e) free-recall, problem solving tasks (integration of part morphology, part physiology, and system operation). Each of the tests was scored on a twenty point scale.

Test One (T1): Free Recall of Cardiac Part Names. The list learning task of Test T1 required the learner to list the names of the parts of the heart and its phases as recalled from the tutorial. Both the verbal and graphic forms of test groups completed the same version of this test.


VALVE CLOSES
VENTRICLE CONTRACTION

The first effect of the pressure produced in the right ventricle is to force the blood behind the flaps of the tricuspid valve, closing the passage-way between the right auricle and right ventricle.

FIGURE 3. Sample information screen from the MCBL

Test Two (T2): Cued Recall of Cardiac Part Identification. Test T2 was a cued-recall measure of the identification of part locations in the heart. Twenty multiple choice items tested the proper identification of a part name in a line drawing of a cross section of the human heart (comparable to ones used in the microcomputer-based instruction). For the verbal form of the test, subjects were asked to identify the correct name (from a list of five) corresponding to a numbered arrow indicating a specific location in the drawing. In the graphic form of the test, subjects were asked to identify which of five labeled arrows designated a specified part name.

Test Three (T3): Cued Recall of Cardiac Part Description. Test T3 represented a simple concept learning task in which learners were required to identify a part by name after being given a description of its structure or function. In the verbal form, this involved selecting the correct name from a list of five to complete a simple statement describing the structure or function of an individual part. The graphic form, called for the selection of a labeled arrow to match the statement about a part.

Test Four (T4): Cued Recall of Cardiac System Functions. Test T4, the cued-recall of system functions, demanded that the learner make responses about relationships among parts of the heart and the cycling of the heart beat. Twenty multiple-choice items involved the complex concept learning task of identifying the appropriate part name(s) or term(s) which best completed a statement about the functions within the human heart. The graphic form of the test accomplished this by asking the learner to choose (from among four drawings of the heart) the one which best depicted the situation described in the item.

Test Five (T5): Free Recall of Cardiac System Functions. The free recall test of system functions (T5) involved the writing or drawing of cardiac system spatial arrangements and functions. The verbal form required that the learner write an essay describing the structure and function of the human heart as recalled from the computer tutorial. This was to include names and location of parts as well as a description of the pumping of blood and related phases. The graphic form required the same information, only drawn in picture form by the learner. Labels and arrows showing parts and blood flow were encouraged.

For the delayed testing session, identical forms of tests T1 through T5 were administered.

Test score results were analyzed through analysis of variance with repeated measures on a university mainframe computer using the SAS© statistical package. For post hoc analysis of multiple comparisons, the Student-Newman-Keuls Test was employed.

Results

The descriptive statistics for performances measured by the five criterion measures are shown in Tables 1 and 2 (immediate and delayed testing sessions, respectively). Graphical representations of these data are depicted in Figures 4 through 9.

From the ANOVAs conducted on the data, the following measurements of significance were made:

1. The animated visual treatment produced higher scores than the static visual treatment with no imagery cueing ($p<.05$, $F=3.37$, $df=2$) on the cued-recall test of system functions when administered immediately after treatment. (See Tables 3 and 4.)

TABLE 1
Descriptive Statistics for Immediate Testing Results

	Learning Tasks									
	T1 List of Parts Free Recall		T2 Part Ident. Cued Recall		T3 Part Descript. Cued Recall		T4 System Fns Cued Recall		T5 System Fns Free Recall	
Treatment Groups	M	SD	M	SD	M	SD	M	SD	M	SD
Static Visual, No Cueing										
Graphic Form of Test	11.83	4.43	13.42	3.29	12.75	3.17	11.42	4.06	6.83	2.35
Verbal Form of Test	13.42	3.37	11.67	4.68	10.00	3.98	9.42	2.78	8.00	2.73
Both Forms	12.63	3.93	12.54	4.05	11.38	3.79	10.42	3.55	7.42	2.56
Static Visual, Imagery Cue										
Graphic Form of Test	13.75	4.00	12.67	3.87	10.92	5.45	12.75	4.03	8.04	4.61
Verbal Form of Test	14.33	3.23	12.17	4.47	11.92	4.03	12.25	2.83	10.33	3.93
Both Forms	14.04	3.57	12.42	4.10	11.42	4.72	12.50	3.41	9.19	4.35
Animated Visual, Imagery Cue										
Graphic Form of Test	15.50	4.01	14.83	2.98	12.83	4.61	12.50	4.10	10.50	3.04
Verbal Form of Test	14.92	3.40	14.67	3.75	13.92	3.53	13.33	3.39	13.88	4.46
Both Forms	15.21	3.65	14.75	3.31	13.38	4.05	12.92	3.71	12.19	4.11
All Treatments	13.96	3.82	13.24	3.93	12.06	4.25	11.94	3.68	9.60	4.20

Note: Maximum score = 20. n = 72.

TABLE 2
Descriptive Statistics for Delayed Testing Results

	Learning Tasks									
	T1 List of Parts Free Recall		T2 Part Ident. Cued Recall		T3 Part Descript. Cued Recall		T4 System Fns Cued Recall		T5 System Fns Free Recall	
Treatment Groups	M	SD	M	SD	M	SD	M	SD	M	SD
Static Visual, No Cueing										
Graphic Form of Test	8.33	4.48	12.17	4.43	12.08	3.70	10.42	3.96	8.71	3.07
Verbal Form of Test	11.42	4.52	9.67	5.07	7.67	3.98	8.42	3.45	5.58	2.53
Both Forms	9.88	4.67	10.92	4.83	9.88	4.39	9.42	3.78	7.15	3.18
Static Visual, Imagery Cue										
Graphic Form of Test	10.00	6.45	11.58	4.50	11.17	4.73	10.75	4.86	7.21	5.01
Verbal Form of Test	14.00	5.08	10.92	4.72	10.75	4.58	11.08	3.45	8.33	4.47
Both Forms	12.00	6.04	11.25	4.52	10.96	4.55	10.92	4.13	7.77	4.68
Animated Visual, Imagery Cue										
Graphic Form of Test	13.00	5.48	14.08	2.84	12.92	3.65	12.25	3.44	8.83	2.85
Verbal Form of Test	12.17	4.73	12.83	3.79	11.75	4.11	12.83	2.98	11.54	5.07
Both Forms	12.58	5.02	13.46	3.34	12.33	3.85	12.54	3.16	10.19	4.25
All Treatments	11.49	5.33	11.88	4.37	11.06	4.33	10.96	3.88	8.37	4.24

Note: Maximum score = 20. n = 72.

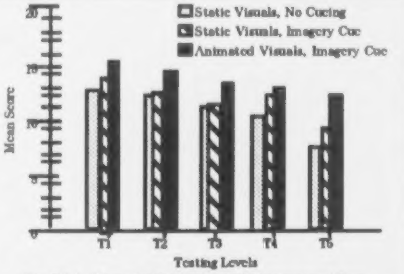


FIGURE 4. Means for Testing Levels by Instructional Type (Immediate Testing)

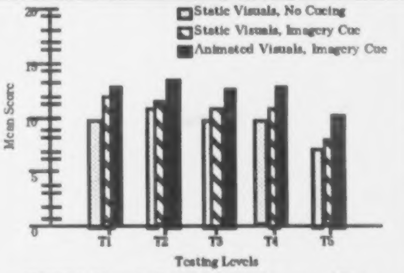


FIGURE 5. Means for Testing Levels by Instructional Type (Delayed Testing)

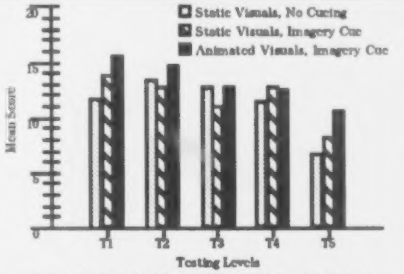


FIGURE 6. Means for Testing Levels by Instructional Type (Immediate Testing, Graphic Form of Test)

2. The static visual treatment with imagery cueing produced higher scores than the static visual treatment with no imagery cueing (p<.05, F=3.37, df=2) on the cued-recall test of system functions when administered immediately after treatment. (See Tables 3 and 4.)

3. The animated visual treatment produced higher scores than the static visual treatment with no imagery cueing (p<.05, F=4.09, df=2) on the cued-recall test of system functions when administered one week after treatment. (See Tables 3 and 5.)

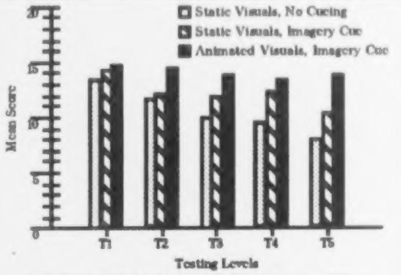


FIGURE 7. Means for Testing Levels by Instructional Type (Immediate Testing, Verbal Form of Test)

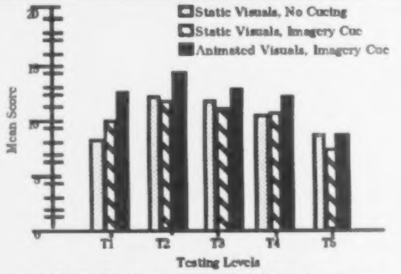


FIGURE 8. Means for Testing Levels by Instructional Type (Delayed Testing, Graphic Form of Test)

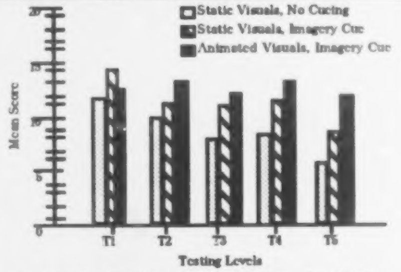


FIGURE 9. Means for Testing Levels by Instructional Type (Delayed Testing, Verbal Form of Test)

4. The animated visual treatment yielded higher scores than each of the static visual treatments with and without imagery cueing (p<.001, F=10.64, df=2) on the free-recall test of system functions when administered immediately after treatment. (See Tables 6 and 7.)

5. The static visual treatment with imagery cueing yielded higher scores than the static visual treatment without imagery cueing (p<.001, F=10.64, df=2) on the free-recall test of system functions when administered immediately after treatment. (See Tables 6 and 7.)

6. Due to an interaction between instructional method and form of test (p<.05) for the delayed testing of the free-recall test of system functions, follow-up analyses were made by individual form of test. (See Table 8.)

7. For the verbal form of the free-recall test of system functions, the animated visual treatment yielded higher scores than the static visual treatment with no imagery cueing (p<.01, F=6.15, df=2) when the test was administered one week after treatment. (See Table 9.)

Table 10 summarizes these results.

Discussion

In this study, the primary focus was on learning higher order concepts (namely dynamic processes involving changes in multiple variables over time) aided by the animation of graphic, instructional elements. Dynamic process was exemplified in the content (delivered by the instructional treatments) as (a) the simultaneous functioning of several parts of the human heart, (b) relationships among parts, and (c) the resulting overall function of the cardiac system through the consequent heart beat cycle.

To indicate the learning effects of animated visuals in the presentation, significant measures in the fourth and fifth tests (cued- and free-recall of cardiac system functions) were sought.

As expected, those tests which did not demonstrate learning of dynamic processes (tests T1 through T3 – list learning of part name, part location identification, and part description) did not show any significant differences among the treatment groups for variations in the instructional method (i.e., visual form or cueing strategy). However, where it was anticipated that the inclusion of animated information would be most effective (tests T4 and T5 – tests of cardiac system functions), significant differences were measured.

On both tests of system operations (cued and free recall), when tested immediately after treatment exposure, learners who experienced the microcomputer-based instruction with animated graphics outperformed those seeing only the static form and no imagery cue strategy.

In fact, both groups who experienced the imagery cue (animated and static graphics) significantly outperformed those who viewed the static visuals, no cue treatment form in the immediate testing session. This leaves some uncertainty as to whether the measured differences should be attributed to the inclusion of the animation or the presence of the embedded cueing strategy.

Two patterns of results, however, may suggest a clearer contribution by the form of visuals.

It was observed that the animated-visual group maintained superiority of performance over the static-visual, no-cue group following the one week delay, while the static-visual, imagery-cue group could not. This pattern of retention was consistent in both formats (graphic and verbal) of the cued-recall test as well as in the verbal format of the free-recall test. This more strongly suggests the role of the animated visual in learning dynamic processes. It appears that subjects in the group experiencing the animated visual treatment were able to better retain information about how the cardiac system operates. Those viewing only static visuals could not be differentiated by the imagery cueing technique following the delay between instruction and testing.

The most pronounced contrast in this study occurred with the immediate

TABLE 3.
Analysis of Variance of
Performance on the Cued Recall of System Functions
by Instructional Methods, Form of Test, and Time of Test

Source	df	SS	MS	F
Between Ss	71	1798.1597		
Instructional Method (A)	2	194.5972	97.2986	4.14*
Test Form (B)	1	7.5625	7.5625	0.32
Method X Form (AB)	2	46.5417	23.2708	0.99
Error Between (S/AB)	66	1549.4583	23.4766	
Within Ss	72	263.5000		
Time of Test (C)	1	35.0069	35.0069	10.62**
Method X Time (AC)	2	8.7639	4.3819	1.33
Form X Time (BC)	1	0.3403	0.3403	0.10
Method X Form X Time (ABC)	2	1.9306	0.9653	0.29
Error Within (SC/AB)	66	217.4583	3.2948	
Total	143	2061.6597		

* p < .05
** p < .01
*** p < .001

TABLE 4.
Analysis of Variance of
Performance on the Cued Recall of System Functions
by Instructional Methods and Form of Test (Immediate Testing)

Source	df	SS	MS	F
Instructional Method (A)	2	86.1111	43.0556	3.37*
Test Form (B)	1	5.5556	5.5556	0.43
Method X Form (AB)	2	24.1111	12.0556	0.94
Error (S/AB)	66	844.0000	12.7879	
Total	71	959.7778		

* p < .05
** p < .01
*** p < .001

TABLE 5.
Analysis of Variance of
Performance on the Cued Recall of System Functions
by Instructional Methods and Form of Test (Delayed Testing)

Source	df	SS	MS	F
Instructional Method (A)	2	117.2500	58.6250	4.19*
Test Form (B)	1	2.3472	2.3472	0.17
Method X Form (AB)	2	24.3611	12.1806	0.87
Error (S/AB)	66	922.9167	13.9836	
Total	71	1066.8750		

* p < .05
** p < .01
*** p < .001

TABLE 6.
Analysis of Variance of
Performance on the Free Recall of System Functions
by Instructional Methods, Form of Test, and Time of Test

Source	df	SS	MS	F
Between Ss	71	2285.0815		
Instructional Method (A)	2	384.4618	192.2309	7.28**
Test Form (B)	1	56.8767	56.8767	2.15
Method X Form (AB)	2	100.6701	50.3351	1.91
Error Between (S/AB)	66	1743.0729	26.4102	
Within Ss	72	297.6250		
Time of Test (C)	1	54.3906	54.3906	21.89***
Method X Time (AC)	2	18.5729	9.2865	3.74*
Form X Time (BC)	1	37.5156	37.5156	15.10***
Method X Form X Time (ABC)	2	23.1563	11.5781	4.66*
Error Within (SC/AB)	66	163.9896		
Total	143	2582.7065		

* p < .05
** p < .01
*** p < .001

TABLE 7.
Analysis of Variance of
Performance on the Free Recall of System Functions
by Instructional Methods and Form of Test (Immediate Testing)

Source	df	SS	MS	F
Instructional Method (A)	2	279.1736	139.5868	10.64***
Test Form (B)	1	93.3889	93.3889	7.12**
Method X Form (AB)	2	14.6319	7.3160	0.56
Error (S/AB)	66	866.1250	13.1231	
Total	71	1253.3194		

* p < .05
** p < .01
*** p < .001

TABLE 8.
Analysis of Variance of
Performance on the Free Recall of System Functions
by Instructional Methods and Form of Test (Delayed Testing)

Source	df	SS	MS	F
Instructional Method (A)	2	123.8611	61.9306	3.93*
Test Form (B)	1	1.0035	1.0035	0.06
Method X Form (AB)	2	109.1944	54.5972	3.46*
Error (S/AB)	66	1040.9375	15.7718	
Total	71	1274.9965		

* p < .05
** p < .01
*** p < .001

testing of the free recall of system functions. On this level, the performances of learners who viewed the animated presentation were significantly higher than those of learners who saw the static form, both with and without the imagery cue. In this case, learners appear to have been helped by seeing the heart animated for the free-recall task, when performed immediately following instruction.

Conclusion

The results of this study show that the use of animation in microcomputer-based instruction to teach a dynamic process can be beneficial. While the present study did not clearly distinguish the impact made by the animation of visuals from that made by the presence of an external, mental imaging strategy in all aspects, certain

BEST COPY AVAILABLE

TABLE 9.
Analysis of Variance of
Performance on the Delayed Testing of Free Recall of System Functions
by Instructional Methods for Graphic and Verbal Forms of Test

Source	df	SS	MS	F
Graphic Form of Test	35	488.2500		
Instructional Method (A)	2	19.6250	9.8125	0.69
Error (S/A)	33	468.6250	14.2008	
Verbal Form of Test	35	785.7431		
Instructional Method (A)	2	213.4306	106.7153	6.15**
Error (S/A)	33	572.3125	17.3428	

* $p < .05$
** $p < .01$
*** $p < .001$

results were clearly supportive of the contribution to learning made by animating the visuals. Maintained significance in the retention measures and association of significance to the specific learning tasks associated with dynamic processes presented from the content, suggest potential areas of focus for further study.

This study, like others recently reported links the use of animation in microcomputer-based instruction to specific learning outcomes. Such linkage of the use of a visual form to a particular instructional purpose implies the complex relationship between forms of visual information and instructional function.

From an instructional design perspective, it appears that the inclusion of animation adds another facet to the visual qualities of instructional systems. Benefitting from computer-based authoring systems which boast improved memory, processing times, and functionality, instructional designers are likely to exercise this increasing capability to produce greater numbers of animation-enhanced, microcomputer-based instructional systems. As the increased availability of hardware and software systems make the design and use of this visual form accessible, further evidence of the impact of animated visual information on specific learning tasks (e.g., knowledge of dynamic processes) merits further attention and exploration.

TABLE 10.
Summary of Analyses of Variance of
Performance on the Recall of Cardiac System Functions

	Instructional Methods		
Performance Measures	Static Visuals, No Cueing (SVNC) superior to:	Static Visuals, Imagery Cue (SVIC) superior to:	Animated Visuals, Imagery Cue (AVIC) superior to:
Cued Recall of System Functions			
Immediate Testing (H01)	----	SVNC*	SVNC*
Delayed Testing (H02)	----	----	SVNC*
Free Recall of System Functions			
Immediate Testing (H03)	----	SVNC***	SVIC, SVNC***
Delayed Testing (H04)			
Graphic Form of Test	----	----	----
Verbal Form of Test	----	----	SVNC**

* $p < .05$
** $p < .01$
*** $p < .001$

References

Baek, Y.K., & Layne, B.H. (1988). Color, graphics, and animation in a computer-assisted learning tutorial lesson. *Journal of Computer-Based Instruction, 15(4).*

Belland, J.C., Mayton, G.B., Dwyer, F.M., & Taylor, W.D. (1991). *Effects of embedded instructional strategies on learning from visuals in microcomputer-based instruction.* Paper presented at the annual conference of the Association of Educational Communications and Technology, Orlando, FL.

Belland, J.C., Taylor, W.D., Canelos, J. Dwyer, F.M. & Baker, P.R. (1985). Is the self-paced instructional program, via microcomputer-based instruction, an effective method of addressing individual learning differences. *Educational Communications and Technology Journal. 33* (3).

Canelos, J., Dwyer, F.M., Taylor, W.D., Belland, J.C., & Baker P.R. (1989). The effect of embedded learning strategies in microcomputer-based instruction. *Journal of Experimental Education, 57* (4).

Canelos J., Taylor W.D., Dwyer F.M & Belland, J.C. (1988). Cognitive style factors and learning from microcomputer-based instructional materials. In M. Simonson (Ed.), *Proceedings of the Research & Theory Division—Association for Educational Communications & Technology.* Ames, IA: Iowa State University.

Dwyer, C.A. (1984). The effect of varied rehearsal strategies in facilitating achievement of different educational objectives as measured by verbal and visual testing modes (Doctoral dissertation, The Pennsylvania State University, 1984). *Dissertation Abstracts International, 45*, 10A3059.

Dwyer, F.M. (1972). *A guide for improving visualized instruction.* State College, PA: Learning Services.

Dwyer, F.M. (1978). *Strategies for improving visualized instruction.* State College, PA: Learning Services.

Rieber, L.P. (1988). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. (ERIC document Reproduction Service No. ED 295 659)

Taylor, W.D., Canelos, J., Belland, J.C. Dwyer, F.M. & Baker, P.R. (1987). The effects of program embedded learning strategies, using an imagery cue strategy and an attention directing strategy to improve learning from microcomputer based instruction. In M. Simonson (Ed.), *Proceedings of the Research and Theory Division—Association for Educational Communications & Technology.* Ames, IA: Iowa State University.

Title:

Instructional Analogies and the Learning of Tangible and Intangible Concepts

Authors:

Timothy J. Newby
Donald A. Stepich

Abstract

The effect of instructional analogy training on the level of immediate, as well as 14-day delay, comprehension of tangible and intangible physiological concepts was investigated. Ninety four college-aged subjects were given training either with or without instructional analogies over five tangible and five intangible advanced physiological concepts. Results showed significantly higher scores of comprehension, both immediately and after the delayed period, for those subjects who had received the analogies. Moreover, subjects receiving analogies reported higher perceived levels of lesson enjoyment. Results are discussed in terms of the prescriptive use of analogies within instructional materials and on future research possibilities.

Instructional Analogies and the Learning of Tangible and Intangible Concepts

An instructional analogy has been defined as an explicit, nonliteral comparison between two objects, or sets of objects, that describes their structural, functional, and/or causal similarities (Stepich & Newby, 1988b). An example includes: A red blood cell is like a truck in that they both transport essential supplies from one place to another through a system of passageways.

Commonly utilized as instructional tools (e.g., Curtis & Reigeluth, 1984), analogies have been employed to teach a variety of subjects, including science (Cavese, 1976; Last, 1985; Scheintaub, 1987), computer programming (Rumelhart & Norman, 1981), composition (Ledger, 1977), and creative problem solving (Gordon, 1961). In a particularly innovative application, Nichter and Nichter (1986) taught rural villagers in India principles of health and nutrition by likening them to more familiar principles related to planting and tending crops. Their purpose is to allow relational information to be mapped from a source known to the learner to one that is unknown (Vosniadou & Schommer, 1988).

Instructional analogies have been shown to consist of four basic components: (a) the target domain (or subject); (b) the base domain (or analog); (c) the connector; and (d) the ground (ref.). The target domain refers to the new to-be-learned information. From the previous example, the target would be the red blood cell. The base domain (truck, from the example) consists of information familiar to the learner which will be used to make a comparison. The connector is a verb phrase, such as is like, which establishes the nature of the relationship between the base and target domains (Rumelhart & Norman, 1981). Finally, the ground is a detailed description of the similarities, and possible differences, indicated by the connector. It is represented by the phrase, transport essential supplies from one place to another through a system of passageways.

Theories of analogical transfer have been developed to explain how information from a base domain is used to facilitate the understanding or manipulation of information in another unrelated target area (Gentner, 1982; 1983; 1988; Gentner & Toupin, 1986; Holyoak, 1984; 1985) In most cases, analogical transfer has come to be viewed as a process of "second order modelling" (Holyoak, 1985) in which a model of the base is used to progressively develop a model for the target. This process takes place through the mapping of a limited set of properties between the domains. Central to this conceptualization is that prior knowledge, which is organized and stored in the learner's memory, serves as a framework or "assimilative context" for the acquisition of new knowledge (Glass & Holyoak, 1986; Mayer, 1979).

Although analogies have been frequently utilized and have become integral parts of accepted theories of instructional design (e.g., Reigeluth & Stein, 1983), research to this point has been divided in terms of their effectiveness for the comprehension and retention of concepts. Drugge and Kass (1978), for example, found that verbal analogies did not significantly increase immediate comprehension. Gabel and Sherwood (1980) in a year-long study of analogies within a high school chemistry curriculum demonstrated no significant improvement on chemistry achievement. Likewise, Bean, Singer, and Cowen (1985) found analogies were not effective with above average students.

With these inconsistent findings the limitations of the analogies and their parameters for effective learning should be investigated in order to successfully predict when their use would be beneficial within a particular instructional setting. To date, a number of factors have been identified. The first, and most prominent, is the learners' comprehension of the analogy used to teach the new content. In the Gabel and Sherwood (1980) study, analogies were not helpful to all students. However, it was shown that as many as 48% of the subjects did not fully understand the analogies used to teach the content. Of those that did understand, scores on the semester achievement tests were significantly higher. Similar results were found in later studies designed to identify difficulties in chemistry problem solving (Gabel & Samuel, 1986; Gabel & Sherwood, 1984).

A second limitation of analogies is the tendency to overgeneralize. Overgeneralization refers to the tendency to include things in a category when they, in fact, don't belong. Schustack and Anderson (1979), for example, asked subjects to read and recall brief biographies of fictional characters. When asked to identify statements they had seen before, subjects showed a higher frequency of false recognitions when the fictional biography was closely analogous to the life of a famous real person. In a recent study by Halpern, Hanson, and Riefer (1990) subjects receiving analogies which were designed from base domains significantly different than the target domains proved to be more successful than those derived from similar or near base domain subject matter. Their conclusions indicated that those from the far domain required additional depth of processing to successfully complete the structural mapping and thus the extra effort increased their abilities and subsequent performances. Another interpretation, however, could include that those from the near domain may have had increased numbers of mistakes due to overgeneralization.

Another limiting factor appears to be the time required to make use of analogies. Analogies are effective, but may not always be efficient as instructional aids. In two sets of studies, Simons (1982; 1984) noted that including an analogy in printed instructional materials increased recall and comprehension of newly learned information, but only under conditions of unlimited study time. Restricting the amount of time the subjects were given to read the materials reduced the advantage of the analogy based instruction. According to Simons, analogy based materials require more time because the additional information in the analogy must be read and compared to the other information in the text. This additional effort pays off in subsequent reading of the same materials, however. Learners can often reread text with analogies more rapidly than text without analogies because of the deeper conceptual understanding they gained from the first reading (Simons, 1984).

A final limiting factor is the learner's need for cues indicating the relationship between the information to be learned and its analog. Cueing is particularly important because learners do not always see the relationship between an analog and its target. As a result, they do not always use the analog when performing the target task (Gick & Holyoak, 1980; 1983). Reed, Dempster, and Ettinger (1985) tried a variety of cueing techniques to increase the transfer of information between algebra word problems. These included describing the relevancy of the analog, making the analog solution available while solving the target problem explaining why a particular equation was used to solve an analog problem, and matching the complexity of analog and target problems. Their failure to produce consistent results demonstrates the difficulty learners have in applying analogous information, even when its usefulness is highlighted.

As indicated by the previous studies, the emphasis on the study of analogies and their impact on learning has been focused on the analogy itself (i.e., how they are constructed, cued, and placed within the instructional materials). A second area of research however, should also extend to the type of to-be-learned concept or material within the target domain which is to be taught. Are analogies more or less effective when the to-be-learned concepts vary in difficulty, ambiguity, complexity, or abstractness? Newby and Stepich (1987) for example, have argued that abstract concepts are qualitatively different than concrete concepts in ways that make them more difficult to learn. An analogy, in their view, can facilitate learning an abstract concept by generating a "prototype substitute" that can represent the abstract concept in memory in much the same way that a prototype comes to represent a concrete concept in a concept learning task. Davidson (1979) has also described analogies as a way of translating abstract information into a form that is more concrete and imaginable and, therefore, more easily understood. According to Simons (1982; 1984) this is the "concretizing" function of analogies.

The concretizing function was demonstrated in a lesson designed by Iona (1982) to teach college students about electricity. The more abstract components of an electrical system were likened to the more concrete and imaginable components of a hydraulic system in which water flows from a hilltop reservoir to a mill at the bottom of the hill. For example, in the analogy electrical voltage was likened to the distance between the reservoir and mill; amperage was likened to the rate of water flow; and electrical resistance was likened to narrow pipes or anything else that will obstruct the flow of water. Other concretizing applications can be found in subjects as diverse as biology (Cavese, 1976) and political science (Russell, 1980).

The present study was designed to compare the effectiveness of instructional analogies given different types of concepts within a single target domain. Specifically, five concepts rated as highly tangible and five rated as highly intangible, all from the same content area of physiology, were selected for this study. Two groups of subjects received training involving all concepts. Training for one group included instructional analogies for each of the concepts; whereas, the training for the other group omitted the analogies. The main purpose of the investigation was two-fold: (a) to investigate if subjects given instructions with analogies comprehend and recall concepts more effectively than subjects who receive instructions without the analogies; and (b) to study the differential effects of analogies given tangible and intangible physiological concepts.

The design of this study also allowed for the investigation of several additional questions involving the use of analogies. For example, "Would an intervening period of time between training and additional testing affect the group performances?"; moreover, "Would there be a reported change in the degree of comprehension and recall based on concept type?"; "Would using analogies result in greater confidence in learning?"; "Would using analogies result in greater lesson enjoyment?"; and "Would the times required for initial learning and/or testing of the materials differ between the groups?".

Method

Subjects

Ninety four subjects (72 female, 22 male) from an undergraduate introductory educational psychology course at a major midwestern university were solicited to participate in the study. All subjects volunteered in order to meet a course requirement for participation in research. The participants' declared major fields of study included education (64.9%), humanities (14.9%) science (11.7%), and physical education (3.2%). Eight subjects failed to return for the two-week follow-up testing. These included six who had been assigned to the analogy group and two from the no-analogy group.

Instructional materials

Concept selection. Physiology was selected as the content area for the study based on the premise that it is a concept-rich subject that includes a range of concepts from the very tangible to the very intangible. Additionally, this subject matter was predicted to be highly unfamiliar to the target group of subjects. A physiology instructor from the Purdue University School of Veterinary Science and Medicine and two of his graduate assistants served as content experts for the development of the materials.

The content experts used their course textbook, Physiology: A Regulatory Systems Approach (Strand, 1983), to identify 611 potential target concepts. This list was reduced to 32 using the following rules: (a) concepts that labeled structures were eliminated while concepts that labelled processes were kept; (b) two word concepts were eliminated while one word concepts were kept; (c) concepts common in everyday language (e.g. salivation, respiration) were eliminated.

In order to make the final selection of concepts from this list of candidates, the 32 concepts were rated by 10 experts (professors, instructors, and advanced graduate students) from Purdue University's School of Veterinary Science and Medicine and the School of Science. Using a seven-point semantic differential rating instrument, each expert rated each of the 32 concepts (Kerlinger, 1973). Five concepts rated as the most tangible (ossification, parturition, micturition, adaptation, and peristalsis) and five rated as the most intangible (disinhibition, pinocytosis, adsorption, summation, and catabolism) were selected for the study.

Instructional materials development. The purpose of the development phase was to create the lesson, in analogy and no-analogy versions, that would be used to teach the concepts selected in the preceding stage. In its final form, the lesson included the following written materials: (a) an introduction and instructions; (b) instructional materials for each concept, including a definition, a one-paragraph description of the physiological process, a verbal analogy (for the analogy condition only), a posttest, and a follow-up questionnaire. The development and validation of the materials followed accepted instructional design principles (e.g., Dick & Carey, 1985).

The first step in creating the lesson was to meet with the content experts and construct an analogy for each of the selected concepts. Construction of the analogies followed the steps outlined by Stepich and Newby (1988b). For each concept, the feature most important to comprehension was identified and one or more concrete items having the same or a similar feature were listed. One of these concrete items, likely to be familiar to the learners, was then chosen as an analog. The analogy was completed by describing the similarities between the chosen analog and the concept. As an illustrative example, the feature most important to understanding the process of peristalsis is the progressive wave of muscular contraction propelling food through the digestive tract. Potential analogs included extracting toothpaste from a tube or squeezing ketchup out of a single-serving packet. Both analogs were expected to be familiar to the learners and the ketchup squeezing analog was chosen as the more accurate of the two. Peristalsis was then described in terms of the analog:

> "Peristalsis is like squeezing ketchup out of a single-serving packet. You squeeze the packet near one corner and run your fingers along the length of the packet toward an opening at the other corner. When you do this, you push the ketchup through the packet, in one direction, ahead of your fingers until it comes out of the opening."

This process was repeated for each of the other nine concepts. Each of the analog or base domains were selected based on two criteria: (a) their high degree of familiarity for the learners (Gabel and Samuel, 1986) and (b) they were from a different or far domain than the to-be-learned concepts (Halpern, et al., 1990).

Next, several physiology textbooks (Holmes, 1979; Jacob & Francone, 1965; Luciano, Vanden, & Sherman, 1983; Parker, 1984; Schmidt & Thews, 1983; Strand, 1983; Vanden, Sherman, & Luciano, 1983) were used as information sources and an initial draft of the lesson was written. This draft included an introduction to the study, a definition, description, and analogy for each concept, and a follow-up questionnaire. This draft was then evaluated by the content experts and two experienced instructional designers to ensure its accuracy, clarity, and appropriateness for the intended subject group. Suggestions obtained were incorporated into a second draft of the materials. The concept materials were again evaluated by the content experts and their suggestions were incorporated into a third draft.

At this point the content experts wrote a set of 20 multiple-choice test items (two items per concept). Each item was developed to focus on concept application, as opposed to simple recall. All test items were then ordered randomly. The test directions and items were then reviewed by an instructional designer and experienced teacher who had not previously seen the materials and who was not a content expert. This helped to ensure that the tests were clear and comprehensible. Evaluative comments were incorporated into a revised version of the test.

The follow-up questionnaire was next constructed. This consisted of questions to obtain bibliographic information (e.g., sex, age, academic major), a question asking the participant to estimate how many questions on the 20-item test they answered correctly, and a Likert-type rating scale to indicate their degree of enjoyment with the lesson.

The development of the materials was followed by a field test involving 24 subjects. Each was given a complete set of materials which included an introduction and set of instructional materials, a posttest, and a follow-up questionnaire. The instructional materials differed based on the independent variable; however, tests and questionnaires were identical for all subjects. Reliability scores, using the Kuder-Richardson formula (Mehrens & Lehman, 1984) indicated a posttest reliability of .68. Following suggestions given by the participants, the wording of several questions in the questionnaire was revised and the verbal instructions given to introduce the study were incorporated within the written instructions.

Procedures

For the formal investigation, subjects were allowed to sign up for one of three scheduled 90-minute periods. Each session was scheduled in a university classroom that could facilitate up to 50 students. Investigators monitored each experimental session and were given a set of procedures to follow. A digital clock was placed at the front of the room to provide consistent time to be recorded. When the subjects were seated the investigator briefly introduced the study and distributed the handout containing the introduction and concept lesson. This handout had been stacked alternating between analogy and no-analogy versions. The copies were then distributed by rows, effectively randomizing assignment to the two experimental conditions.

The concept lesson asked the subjects to note the last four digits of their social security number (for identification purposes) and to record the time they began studying the materials. After they had studied the concept materials for as long as they wished, they were again asked to note the time. The first handout was then returned to the investigator whereupon a copy of the posttest was received. All subjects received the same posttest. Subjects were asked to record the 4 digit identification code, as well as the starting and completion times of the test. No limits were given on the amount of time to finish the exam.

After the test was completed, subjects returned it to the instructor and they were given a copy of the follow-up questionnaire. Upon completion and return of the questionnaire each subject was reminded to return in 14 days at the same time and location. No further instructions were given and they were free to leave.

When the subjects returned in 14 days they were each given a second posttest. This test consisted of a short set of directions and the same 20 questions (in a new random order) of the initial posttest. Subjects were also asked to record their four-digit social security code number and the beginning and ending time of the test. No time limit was given for the test and it was returned to the investigator upon completion. All subjects were then debriefed and thanked for their participation.

Results

This investigation examined four dependent variables: comprehension of the concepts (both immediate and longer term); enjoyment of the lesson; confidence in learning; and time required to study the concepts and complete the posttests. In each case separate analyses were completed.

To compare the comprehension of the concepts across conditions, the 20-question immediate posttest was graded for each subject and the results were grouped based on method of instruction and type of concept. A mixed-factorial analysis of variance (method of instruction by type of concept with repeated measures) was performed on the mean number of items answered correctly for each concept. As shown in Figure 1, those subjects receiving instructions with analogies significantly outperformed those who did not receive the analogies (F (1,92) = 10.53; MS = 67.92; $p < .002$). No significant difference was shown between the comprehension of the tangible and intangible concepts; however, a significant interaction (F (1,92) = 6.09; MS = 17.28; $p < .02$) between the two independent variables was recorded. Figure 1 illustrates that those subjects in the analogy group performed at a slightly higher level given tangible concepts than they did given intangible concepts; whereas, those in the no-analogy group achieved a higher score given intangible concepts when compared with their scores for the tangible concepts.

Insert Figure 1 about here

The two-week posttest was examined in the same manner as that of the initial posttest. As shown in Figure 2, the two groups attained the same mean scores as reported for the initial posttest for the tangible concepts but decreased in their comprehension for the intangible concepts. A significant difference was again recorded based on the method of instruction as those receiving analogies significantly outperformed those not receiving such instructions ($\underline{F}$ (1,84) = 9.43; $\underline{MS}$ = 67.37; $\underline{p} < .003$). In this case however, there was a significant difference between the tangible and intangible concepts ($\underline{F}$ (1,84) = 7.17; $\underline{MS}$ = 22.349; $\underline{p} < .009$), but no significant interaction was reported.

Insert Figure 2 about here

Enjoyment of the lesson was measured by a single questionnaire item in which the subjects were asked to rate how much they had enjoyed the lesson on a scale from 1 (not at all) to 5 (a lot). A significant difference was found between the analogy ($\underline{M}$ = 3.02) and the no analogy ($\underline{M}$ = 2.59) conditions using a two-tailed t-test ($\underline{t}$ (92) = 1.869; $\underline{p} <.05$).

Confidence in learning was measured by asking the subjects to predict how many items they felt they had answered correctly on the initial posttest. Means for the analogy and no-analogy conditions were 14.19 and 13.55 respectively. A two-tailed t-test indicated no significant difference between the analogy/no-analogy predictions.

In order to measure the final variable, time, all subjects were asked to record the time they started and finished both the lesson and the two posttests. The time spent, in minutes, was then tallied for the analogy and no-analogy conditions. Two-tailed t-tests were performed on the mean study times and the mean test times. A significant difference was found in the mean study times ($\underline{t}$ (92) = 1.953; $\underline{p} < .05$) but not for either of the testing periods. The mean study times for the analogy and no-analogy conditions were 12.04 minutes and 10.55 minutes, respectively.

Discussion

From the measurement of the four dependent variables, comprehension, enjoyment, confidence, and time, several important findings should now be discussed. First, the results indicate that analogies had a beneficial effect on the comprehension of unfamiliar concepts. In other words, subjects who received instructions which included analogies scored significantly higher on the immediate posttest than those who did not receive training which included the analogies. Moreover, this difference in comprehension was sustained during the two-week posttest. Even though the learners were not prompted to recall or use the analogies in any way, comprehension scores indicate a difference between the two groups remained even after the 14-day interval. The use of analogies appears to be an effective instructional strategy which increases the immediate and long term comprehension of concepts.

Conclusions about the influence of analogies on the comprehension of different types of concepts in the present experiment are strengthened by the fact that the subjects in the analogy condition were neither trained in the use of analogies nor cued to use the analogies they were given. Gick and Holyoak (1980; 1983) and others (e.g., Reed, Dempster, & Ettinger, 1985; Schustack & Anderson, 1979) have shown that training and cueing are essential aspects of using analogies effectively in instruction. The subjects in the present experiment, however, were not given practice in using analogies or cues to use analogies in recalling the concepts for the test. In spite of this, subjects in the analogy condition outperformed subjects in the no-analogy condition on both posttests. This indicates that training and cueing may not be necessary and that analogies facilitate concept learning even when the subjects are neither trained nor cued. Additional research is needed to further explore the contribution training and cueing make to the effectiveness of analogical instruction as well as the best methods for providing such training and cueing.

A second relevant finding from this investigation was the degree of effectiveness of analogies based on the concepts being either tangible or intangible in nature. Although during the immediate posttest the analogy group significantly outperformed the no-analogy group given either type of concept, the analogy group showed no significant difference in the comprehension of the tangible versus that of the intangible concept types. This was not the case, however, after the 14-day interval had elapsed. In that instance, while the analogy group still significantly outperformed the no-analogy group for both types of concepts, the comprehension of the intangible-type concepts decreased significantly for both groups. Although the analogy instruction remained more effective, the concurrent decrease indicates the possibility of interference being greater for those concepts of a less tangible nature. Several authors have pointed out the differences between concept types and the increased difficulty of those of an intangible nature, however, further research focusing on the potential interference differences is needed (Newby & Stepich, 1987; Reed & Dick, 1968).

A final point that should be considered in the discussion of the concept-type differences is that of the interaction effect shown to occur during the initial posttest. As indicated in Figure 1, the no-analogy group initially scored higher for intangible concepts than they reported for the tangible concepts while the opposite was true for the analogy group. Such results by the no-analogy group contradict previous investigations as well as their own 14-day interval posttest results, and thus may be an indication of spurious results (Royer & Cable, 1976; Reed & Dick, 1968). However, this result may also be an indication of the overall effectiveness of the training. For each individual concept the learner received a definition and a description. It may be that those different training lessons for the intangible concepts were initially clearer or more readily understood. These conclusions are questionable however, because of similar effects not being revealed within the data from the analogy group and because of no interaction seen aftger the two-week interval.

A second factor measured during this investigation was enjoyment for the lesson as perceived by the learners. Following training and the initial posttest, all subjects were asked to rate the degree to which they enjoyed the lesson. The results indicate the use of analogies positively influenced the subjects' rated enjoyment of the lesson. That is, analogies may be valuable in instruction because they increase the learners' enjoyment of learning. Keller and Kopp (1987) and Newby (in press) have pointed out that familiarity helps learners see the relevance of what they are learning which increases their motivation to learn. Analogies connect new concepts to familiar objects and events and, thus, may increase the learners' motivation by increasing the relevance of the instruction. The strength of this conclusion is limited, however, because it is based on a single subjective measure. Additional research is needed to explore the effects of using analogies on more objective measures of enjoyment.

The third variable measured the perceived confidence subjects had in their comprehension of the concepts. The results show that using analogies did not affect the subjects' confidence in their learning. This is contrary to what might be expected, especially in light of the observed differences in comprehension between subjects in the analogy and no-analogy conditions. One possible explanation for this result is that the subjects perceived the concepts as so difficult that they saw little chance of success, even with analogies to help them, and so predicted a lower score than they actually obtained. Another possibility is that confidence is not closely related to actual performance. The relationship between analogy and confidence is an important issue because confidence is an important aspect of motivation (Keller, 1983). Additional research is needed to explore the relationship and to determine how analogies might be used to increase the confidence of the learners.

The final measured variable included the amount of time they used to both study the concepts and complete the posttest. The results show that subjects in the analogy condition used significantly more time studying than subjects in the no-analogy condition, but did not require significantly more time to complete the posttest. These results are consistent with the findings of Simons (1982; 1984) in which analogies may be more time-consuming, as well as more effective, than other instructional strategies. It is the efficiency of the instruction rather than its effectiveness that is at issue here. Even when they are equal in length, lessons with analogies require more time than lessons without analogies because the analogies require the learners to process the connection between the target and analog in addition to reading the content. Simons notes that there is a trade-off between the increased time needed to use analogies and their benefits in performance. The results of the present experiment support Simons' conclusion. However, the strength of this support is limited by the fact that the lessons used in the present experiment were unequal in length. For the analogy condition the analogies were simply added to the descriptive paragraphs for each concept, which means that the subjects in the analogy condition had more content to read. Subjects in the analogy condition took more time to study the concepts, but their comprehension was improved. The exact location of the balance point is unclear, however, and may be a purely subjective decision to be made by the individual teacher or learner. Additional research is needed to further investigate the relationship between study time and comprehension in concept learning through analogical instruction.

In conclusion, several lines of reasoning converge to suggest that analogies are potentially powerful instructional tools: (a) their pervasiveness in both everyday and instructional communication; (b) anecdotal evidence of their influence in scientific discoveries throughout history; and (c) empirical evidence of their effectiveness in a variety of learning tasks. However, what has been written about analogies is almost entirely descriptive in nature. There is little prescriptive information and, as a result, few guidelines for using analogies in instructional practice. Newby and Stepich (1987) have taken a step toward filling this gap by suggesting a set of procedural guidelines for both creating and utilizing analogies within instruction. The present investigation was undertaken as an empirical test of one of their prescriptions: that analogies are effective given intangible, as well as tangible concepts. The primary finding confirmed this assumption. Secondarily, the study found that using analogies affects the enjoyment of the learners but not their perceived confidence.

References

Bean, T. W., Singer, H., & Cower, S. (1985). Acquisition of a topic schema in high school biology through an analogical study guide. The National Reading Conference Yearbook, 34, 38-41.

Cavese, J. A. (1976). An analogue for the cell. The American Biology Teacher, 38, 108-109.

Curtis, R. V., & Reigeluth, C. M. (1984). The use of analogies in written text. Instructional Science, 13, 99-117.

Davidson, R. E. (1979). The role of metaphor and analogy in learning. In J. R. Levin & V. L. Allen (Eds.). Cognitive learning in children: Theories and strategies. New York: Academic Press.

Dick, W., & Carey, L. (1985). The systematic design of instruction. Glenview, IL: Scott, Foresman & Co.

Drugge, N. L., & Kass, H. (1978). The effect of selected analogies on understanding of scientific explanations. Abstract of Presented Papers, NARST. Cleveland, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education, The Ohio State University.

Gabel, D. L., & Samuel, K. V. (1986). High school students' ability to solve molarity problems and their analog counterparts. Journal of Research in Science Education, 23, 165-176.

Gabel, D. L., & Sherwood, R. D.. (1984). Analyzing difficulties with mole-concept tasks by using familiar analog tasks. Journal of Research in Science Teaching, 21, 843-851.

Gabel, D. L., & Sherwood, R. D.. (1980). Effect of using analogies on chemistry achievement according to Piagetian level. Science Education, 64, 709-716.

Gentner, D. (1988). analogical inference and analogical access. In A. Prieditis (Ed.), Analogica: Proceedings of the first workshop on analogical reasoning. London: Pittman Publishing Co.

Gentner, D. (1982). Are scientific analogies metaphors? In D. S. Maill (Ed.), Metaphor: Problems and perspectives. Atlantic Highlands, NJ: Humanities Press, Inc.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cognitive Science, 7, 155-170.

Gentner, D., & Gentner, D. R. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner and A. Stevens (Eds.), Mental models. Hillsdale, NJ: Erlbaum.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. Cognitive Science, 10, 277-300.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. Cognitive Psychology, 12, 306-355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. Cognitive Psychology, 15, 1-38.

Glass, A., & Holyoak, K. J. (1986). Cognition (2nd edition). New York: Random House.

Gordon, W. J. J. (1961). Synectics: The development of creative capacity. New York: Harper & Row.

Halpern, D. F., Hansen, C., & Riefer, D. (1990). Analogies as an aid to understanding and memory. Journal of Educational Psychology, 82, 298-305.

Holmes, S. (1979). Henderson's dictionary of biological terms (9th edition). New York: Van Nostrand.

Holyoak, K. J. (1984). Analogical thinking and human intelligence. In R. J. Sternberg (Ed.). Advances in the psychology of human intelligence. Hillsdale, NJ: Lawrence Erlbaum Associates.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), The psychology of learning and motivation. New York: Academic Press.

Iona, M. (1982). Teaching electricity. Science and Children, 19(2), 22-23.

Jacob, S. W., & Rancone, C. A. (1985). Structure and function in man. Philadelphia: W. B. Saunders Company.

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.). Instructional design theories and models: An overview of their current status. Hillsdale, NJ: Erlbaum.

Keller, J. M., & Kopp, T. W. (1987). An application of the ARCS model of motivational design. In C. M. Reigeluth (Ed.), Instructional theories in action: Lessons illustrating selected theories and models. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kerlinger, F. N. (1973). Foundations of behavioral research (2nd edition). New York: Holt, Rinehart and Winston.

Last, A. M. (1985). Doing the dishes: An analogy for use in teaching reaction kinetics. Journal of Chemical Education, 62, 1015-1016.

Ledger, M. (1977). Analogy: A lesson in composition. Exercise Exchange, 21(2), 8-13.

Luciano, D. S., Vanden, A. J., & Sherman, J. H. (1983). Human anatomy and physiology: Structure and function. New York: McGraw-Hill.

Mayer, R. E. (1979). Can advance organizers influence meaningful learning? Review of Educational Research, 49, 371-383.

Mehrens, W. A., & Lehman, I. J. (1984). Measurement and evaluation in education and psychology (3rd edition). New York: Holt, Rinehart, & Winston.

Newby, T. J. (in press). Classroom motivation: Strategies of first-year teachers. Journal of Educational Psychology.

Newby, T. J., & Stepich, D. A. (1987). Learning abstract concepts: The use of analogies as a mediational strategy. Journal of Instructional Development, 10, 20-26.

Nichter, M., & Nichter, M. (1986). Health education by appropriate analogy: Using the familiar to explain the new. Convergence, 19, 63-72.

Parker, S. P. (1984). McGraw-Hill dictionary of scientific and technical terms (3rd edition). New York: McGraw-Hill.

Reed, H. B., & Dick, R. D. (1968). The learning and generalization of abstract and concrete concepts. Journal of Verbal Learning and Verbal Behavior, 7, 486-490.

Reed, S. K., Dempster, A., & Ettinger, M. (1985). Usefulness of analogous solutions for solving algebra word problems. Journal of Experimental Psychology: Learning, Memory, and Cognition, 11, 106-125.

Reigeluth, C. M., & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.). Instructional design theories and models: An overview of their current status. Hillsdale, NJ: Erlbaum.

Royer, J. M., & Cable, G. W. (1976). Illustrations, analogies, and facilitating transfer in prose learning. Journal of Educational Psychology, 68, 205-209.

Rumelhart, D. E., & Norman, D. A. (1981). analogical processes in learning. In J. R. Anderson (Ed.). Cognitive skills and their acquisition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Russell, D. (1980). Political spectrum analogy. Social Education, 44, 519.

Scheintaub, H. (1987). Earth elements by analogy. The Science Teacher, 54(4), 29-30.

Schmidt, R. F., & Thews, G. (1983). Human physiology. Berlin: Springer-Verlag.

Schustack, M., & Anderson, J. R. (1979). Effects of analogy to prior knowledge on memory for new information. Journal of Verbal Learning and Verbal Behavior, 18, 565-583.

Simons, P. R. (1982). Concrete analogies as aids in learning from text. In A. Flammer & W. Kintsch (Eds.), Discourse processing. North Holland Publishing Co.

Simons, P. R. (1984). Instructing with analogies. Journal of Educational Psychology, 76, 513-527.

Stepich, D. A., & Newby, T. J. (1988a). Analogical instruction within the information processing paradigm: Effective means to facilitate learning. Instructional Science, 17, 129-144.

Stepich, D. A., & Newby, T. J. (1988b). Analogizing as an instructional strategy. Performance and Instruction, 27(9), 21-23.

Strand, R. L. (1983). Physiology: A regulatory systems approach (2nd edition). New York: McMillan.

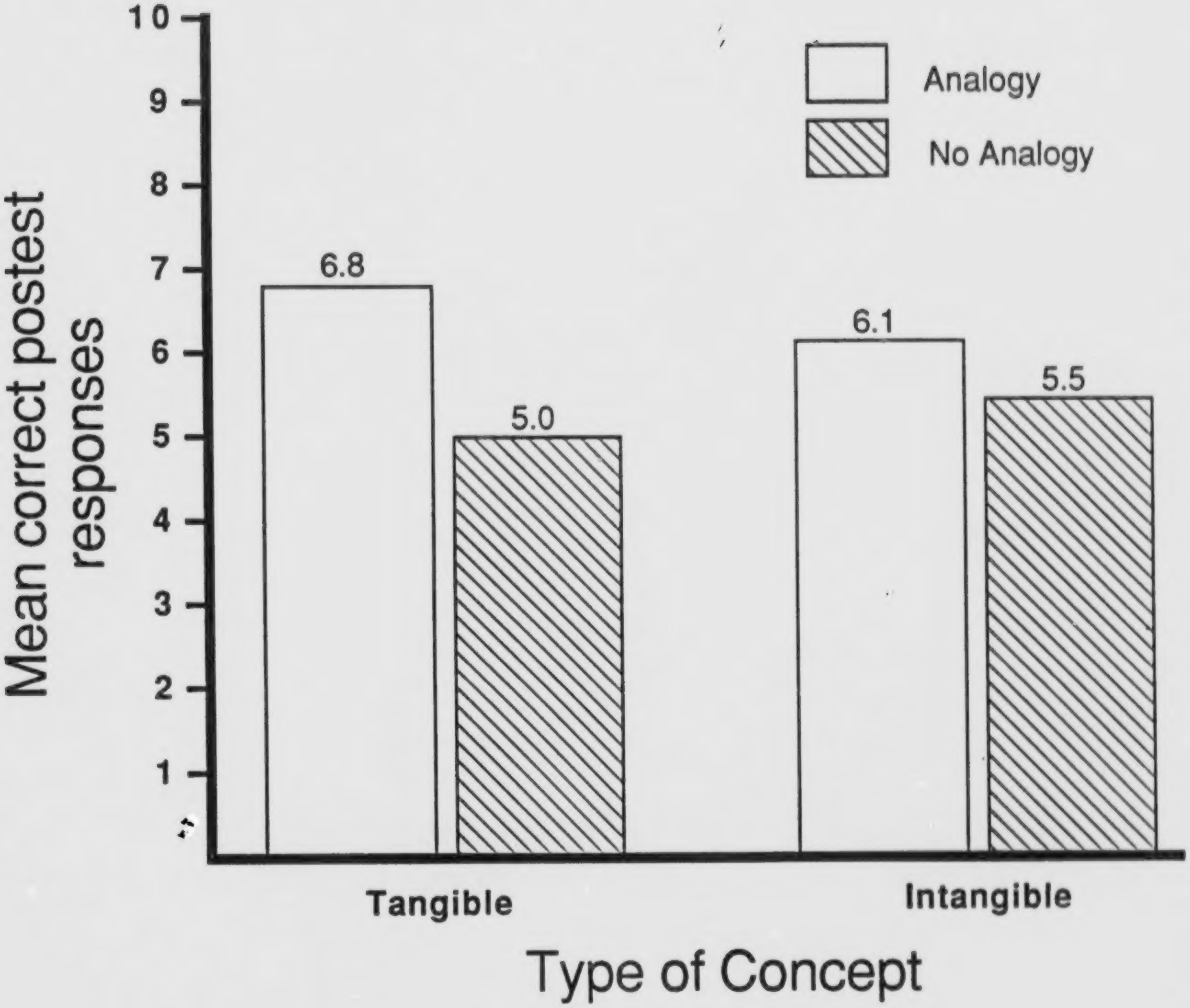
Vanden, A. J., Sherman, J. H., & Luciano, D. S. (1980). Human physiology: The Mechanisms of body function. New York: McGraw-Hill.

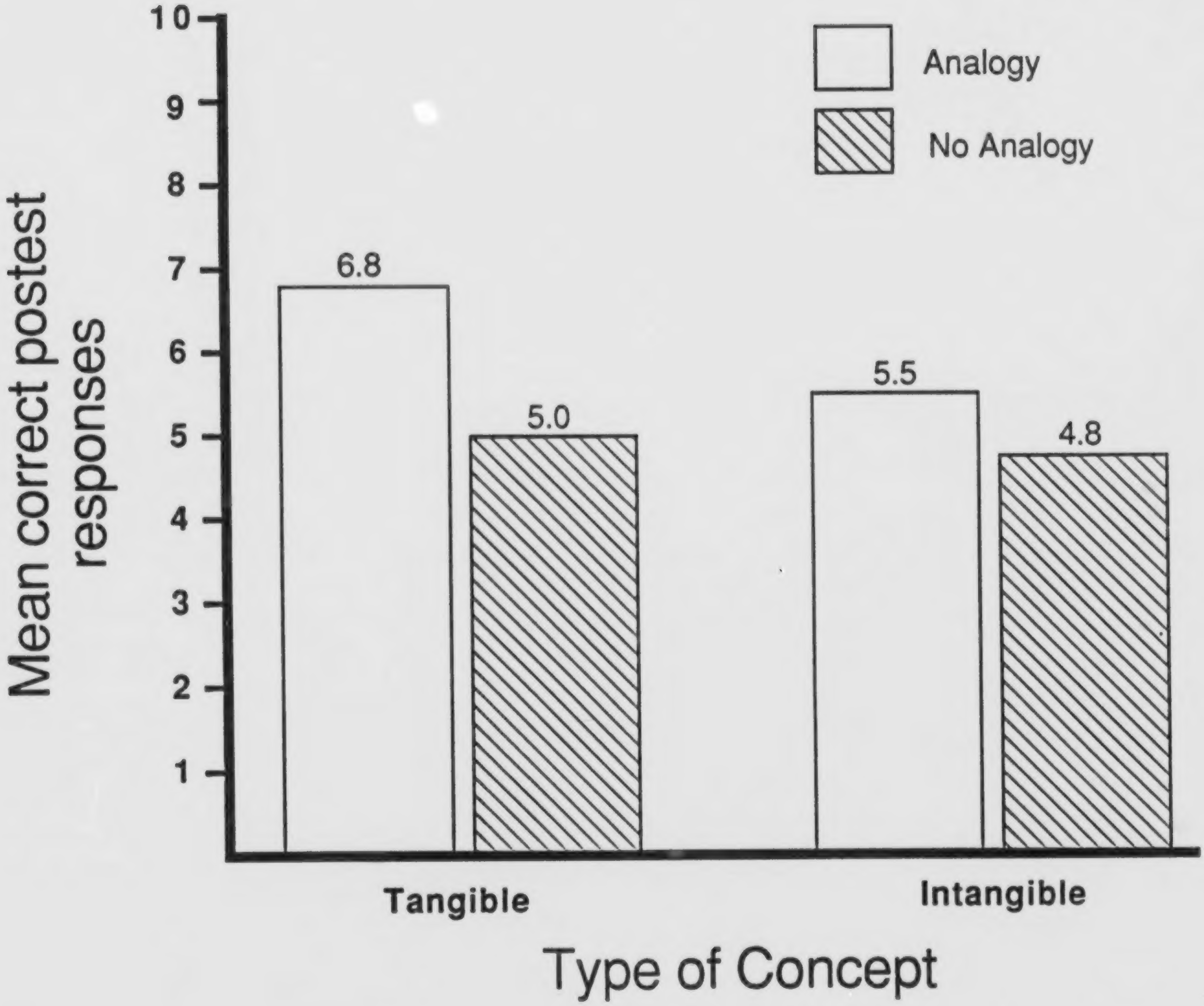
Vosniadou, S., & Schommer, M. (1988). Explanatory analogies can help children acquire information from expository text. Journal of Educational Psychology, 80, 524-536.

Figure Captions

Figure 1. Immediate posttest comprehension mean scores comparing analogy versus no-analogy trained subjects across tangible and intangible concepts.

Figure 2. Fourteen-day posttest comprehension mean scores comparing analogy versus no-analogy trained subjects across tangible and intangible concepts.


10
9
8
7
6
5
4
3
2
1
Mean correct postest responses
Analogy
No Analogy
6.8
5.0
6.1
5.5
Tangible
Intangible
Type of Concept


Mean correct postest responses
10
9
8
7
6
5
4
3
2
1
Analogy
No Analogy
6.8
5.0
5.5
4.8
Tangible
Intangible
Type of Concept

Title:

A National Survey of Systemic School Restructuring Experiences

Authors:

Catherine A. Norris
Charles M. Reigeluth

The idea of restructuring schools, which gained the attention of educators in the middle 1980s and calls for fundamental structural change of our educational system, is now becoming a reality for some schools across the country. These schools, some individually and some with the aid of supporting networks, have undertaken the challenge of reassessing and redesigning the way they think of and do schooling. This article summarizes a project to identify and analyze as many of those schools nationwide as possible.

The School Improvement Resources Inquiry USA project (SIRIUS-A project) is a study funded by the Indiana Department of Education. It has analyzed the accounts of 62 schools' restructuring experiences from across the country. The major purposes of the study are to identify schools throughout the country which are restructuring, to identify the kinds of structural changes being implemented, and to identify the kinds of change processes being used to plan and implement the restructuring. This article will not address the change processes.

Background of the Study

The study is using a case-study approach in order to characterize the uniqueness of each school's restructuring effort and the uniqueness of each school's contextual conditions. Other reasons for using the case study approach are to receive as much information as possible for compiling a holistic description for each school, and to provide a description of the restructuring effort in the school participants' own words.

Because a goal of the study is to identify and survey **all** the schools in the country that are restructuring, the study utilizes a criterion-based sampling strategy.[1] In criterion-based sampling, all cases are included in a study if they meet the established criteria.

For the SIRIUS-A Project, schools which meet the criteria for restructuring are those which have initiated (not just planned) **systemic** restructuring. Systemic means that the changes in the school are interrelated rather than piecemeal, and that change in one part of the school requires changes in other parts of the school. These changes build to a holistic, integrated restructuring effort. The researchers have not included schools that simply used the label of "restructuring."

When restructuring is systemic, all parts of the school are likely to be changed or affected. Examples are the use of "time, talent, and technology,"[2] where time includes periods in the day and grade levels as years, talent includes the roles of teachers, administrators, assistants, and students; and technology includes facilities, equipment, and instructional resources.

One of the participating schools in the study, Linda Vista Elementary School in San Diego, provided an apt definition of systemic restructuring: "School restructuring appears to be most successful if it is a pervasive, systemic change--it affects each student in the school and it does not attack only one aspect of the school program."

Process of the Study

The study began with a search for the population of schools nationwide which may be restructuring. Criteria for systemic restructuring were established, and relevant organizations were contacted. The major sources of contacts for the search were the state departments of education (43 of the DOEs responded). Other major contact sources included the Coalition of

Essential Schools, the National Education Association's Mastery in Learning network, and the North Central Regional Educational Laboratory's Accelerated Schools Action Project. Names of schools were also acquired from the media, current educational publications, and word of mouth from schools participating in the study. The search resulted in an initial data base of 531 schools.

These schools were contacted and given the option of answering an open-ended questionnaire, sending existing information describing their restructuring efforts, or both. Of the 531 schools, 137 (26%) have responded to date (November 1990). The information from the 137 schools was then analyzed to identify examples of systemic restructuring. Those schools which satisfied the criteria of systemic restructuring comprised a final sample of 62 schools, which form the basis of the findings for this article. The remaining schools fell into the categories of "still in the planning stage," "uncertain if systemic, based on their information," or "not systemic, based on their information."

Findings: Not Just Systemic, But Thematic

In the analysis of the 62 schools' information, many schools seemed to have a unified theme in their descriptions of their changes. In contrast, other schools listed changes they were implementing without a readily apparent emphasis connecting the changes. Their nonthematic changes are still systemic in that they appear to affect other changes in the school, are interrelated, and form an integrated system, but they do not seem to be connected by a theme, or at least do not seem to place a strong emphasis on any theme.

For example, the following is a description of a restructuring school that does not appear to emphasize a theme. This school appears to have two general focuses: higher-order learning for students and teacher collaboration. The changes reported by the school include peer tutoring, interdisciplinary team teaching, and presenting parent seminars on thinking skills. The school's changes are systemic in that interdisciplinary team teaching requires teacher collaboration, and peer tutoring and parent seminars on thinking skills are used to facilitate higher-order learning. Yet the portrait of the school does not appear as unified as do the portraits of those schools whose descriptions reveal stronger themes.

While analyzing the schools, it became striking how the systemic descriptions often included an underlying theme explicitly relating the changes. In fact, it appeared that the more a school's changes built upon a theme, the more extensively systemic the restructuring effort seemed to be. The schools with strong themes reported changes that appeared deeper and more interrelated than the changes reported by schools without strong themes.

Thus, perhaps the most significant finding from the SIRIUS-A study is that *some schools seemed to base their restructuring on a connected, underlying theme, which appeared to result in more systemic changes in the school.*

Examples of Theme-Based Restructuring Experiences

The following descriptions of restructuring experiences focus on their systemic and theme-based elements. The descriptions of the schools presented here are not comprehensive; discussion of all the implementations by the schools is not possible in the available space.

Saturn

The Saturn School of Tomorrow in St. Paul, Minnesota, based its design on two major themes: High-Tech, High-Teach, and High-Touch; and mastery learning. The school serves grades 4-7. As an example of a school that explicitly implements its theme, Saturn School reports that its students spend one-third of their time with technology, 1/3 of their time with teachers, and 1/3 of their time working with other students on cooperative learning projects.

The High-Tech component includes the computer-based Integrated Learning System and extensive video-based instruction, especially in reading, writing, and math. The school also uses the Discourse (TM) System for group-based instruction.

High-Teach is the commitment to the belief that students, parents, educators, and the community are instructional resources who can "ensure the success of each student." Some of the ways "High-Teach" is realized are through on-site learning in the community (the St. Paul Public Library, the YMCA, and the Science and Art Museums), through parental involvement, and through a differentiated staff that is not grade- or classroom-level based.

The High-Touch component is described as meeting students with sensitivity and concern on their level. A teaching team stays with a group of students throughout their 3-4 years at the school. All students are also in an advisory group for their full time at the school. In addition, heavy emphasis is placed on letting students follow their interests in what they learn, both in choosing individual activities and in taking heterogeneously-grouped courses which last eight weeks.

Mastery learning is heavily emphasized at Saturn School of Tomorrow. Many of the school's activities are designed in a way that students learn by reaching mastery, rather than by earning grades or passing time in a subject area. Some of these activities are:

- Students and parents develop Personal Growth Plans.
- Students choose much of how they use their time in the school, in order to work toward mastery.
- Mastery learning is monitored through the Integrated Learning System and the Personal Growth Plan process.
- Mastery is assessed through a Portfolio of Proficiencies, a record of teacher comments form courses, Personal Growth Plan documents, and CAI records.

Saturn School's implementations may at first glance seem varied and unconnected, yet the components and the changes within them are interrelated in working toward achieving the High-Tech, High-Teach, High-Touch and mastery themes. For example, the computer-based learning system links a technology component with a mastery component--the technology allows the students to work at their own pace, it monitors the students' attainment of mastery, and it provides records of mastery to be used as assessments.

The differentiated staffing (High-Teach), gives students the individualized time to work toward mastery, and facilitates cooperative learning (High-Touch) in the courses.

Thus, systemic restructuring is demonstrated in Saturn School through the interrelationships among features of the school. These interrelationships permeate the entire school, and present a holistic picture of a fundamentally transformed school. Furthermore, Saturn's two integrating themes of High-Tech, High-Teach, & High-Touch and mastery learning contribute to the extensiveness and depth of Saturn's systemic characteristics.

Skowhegan

Skowhegan Area Middle School of Skowhegan, Maine, has chosen to restructure based on a theme of a "collegial/team approach to change." The school's staff have organized themselves into five teams, each of which decides and implements its own plans each year as schools-within-a-school. Each team creates its own plans for meeting the needs of its students, tries the plans out for a year, and, based on the trial implementations, revises, extends, or drops its various changes. The students with each team are multi-grade and stay with their school-within-a-school throughout their years at the middle school.

Just some examples of implementations within teams are an individualized reading program with daily sustained silent reading for all students, flexible scheduling and multi-grade grouping in one team's math program to meet varying student needs, and a homework monitoring program for "at risk" students. The school also pursues a school-wide action plan, not only to address the needs of all students, but also to address certain goals that all teams need to address at the same time "to provide for a cohesive forward movement."

In order to sustain the multi-grade organization of teams which is Skowhegan's restructuring emphasis, the teams maintain heterogeneous grouping within the classrooms. To address individual student needs at the different grade levels, cooperative learning and differentiated instruction are used in the classrooms. Other key features to enable the team approach are common team planning times, flexible scheduling, peer tutoring, and the use of four para-educators assisting in classrooms to meet mainstreamed students' needs.

Narragansett

Another school in Maine, Narragansett School in Gorham, is an elementary school whose two major themes are for the school to become a center of inquiry and to focus on children's development as learners.

Toward its goal of making the school a center of inquiry, Narragansett has implemented several activities. One of them is developing and carrying out research projects. The school utilizes collaborative decision making. A teacher leader position has been established for each grade to help guide the decision making, and teacher assistants are used to give teachers more release time for peer conferencing. Finally, the school added a teacher-scholar position to observe and facilitate teacher reflection processes.

Narragansett implemented several changes to focus on children's development as learners. Multi-age teacher teams stay with children for more than one year. Parents choose the students' placements with their teachers. Differentiated staffing with teacher assistants is used, and parents volunteer in the classrooms, with some working on teacher-volunteer teams. The students learn with a variety of materials, and are evaluated through alternative forms of assessment, including a multi-media portfolio project. In a second correspondence from Narragansett, we learned that the school is also focusing on helping children in their metacognitive development. In essence, Narragansett has based its restructuring on its two themes by implementing changes that extend directly from them.

Mark Twain

Mark Twain Elementary School in Littleton, Colorado, restructured based on a theme of "The Peak Performance School," coined by the school's principal, Monte Moses, who authored a book with that title.[3] According to Moses, peak performance schools have a clear vision and purpose, seek to actualize human potential, and surpass expectations. He believes that becoming a peak performance school requires first creating a visionary perspective, and then from that vision, establishing a mission for the school.

The vision created by Mark Twain Elementary resulted in a mission to foster human growth. This resulted in three major new features of the school. First, Mark Twain changed its curriculum and assessment from being organized around a body of information to being organized around a small set of tasks and critical knowledge that have utility in a variety of contexts. The restructuring effort established several performance assessments: a Peak Performance Profile (a checklist of character traits, critical thinking skills, reading, writing, and scientific problem solving), a Fifth Grade Research Performance Assessment, and portfolios of projects, which include a self-improvement goal and a service goal. Furthermore, students move at their own pace within grade levels.

Second, the restructuring effort reorganized the teachers into teams that stay with the same students for two or three years. This helps teachers better respond to student needs and differences; as the school contends, students will grow more with teachers who know a lot about them.

Third, Mark Twain now utilizes differentiated staffing arranged in a professional hierarchy of a Lead Teacher, professional teachers, interns, undergraduate aides, and classroom teacher's aides. The larger number of staff gives teachers more time to educate in a professional manner, thus serving the growth needs of students. In addition, there are 15 parent volunteers in the school daily and over 20 more in the school regularly.

Bloomfield Hills

Bloomfield Hills Schools, in Bloomfield Hills, Michigan, is creating a Model High School with extensive changes, through funding from RJR-Nabisco. Its major themes are 1) utilizing community involvement in what the educational process should provide as a basis for designing education, 2) being inquiry-based, 3) using mastery learning, 4) fully integrating disciplines and employing an encompassing theme (such as "How have humans dealt with the question of whether to live in harmony with or dominate nature?"), and 5) giving students a high degree of responsibility and choice.

Students have choice and responsibility for what problems they will study, how to study them, how to use their time to study them, and how they will demonstrate mastery of core competencies in studying them. They decide how they will be assessed for the core competencies, who will assess them (who will serve on their Student Assessment Panel), and even whether they will spend all or a portion of their day at Model High School (with the other part of the day at their former high school in the district).

To create a school that accomplishes these ends, Bloomfield Hills Model High School has implemented major changes in the use of time, in the teachers' role, and in assessment. It has restructured time by offering two 2-hour, interdisciplinary instructional blocks daily, a 25-minute student advisory program daily, and a 2-hour period for students' independent projects (e.g. individual or group, research or internship). The school day also provides one daily period each in common team planning time and in individual planning time for all teachers.

The Model High School restructured the teachers' role in that they are truly co-learners and facilitators with the students. Because the teachers are creating the new thematic curricula based on community input and on their own and their students' explorations of how disciplines interrelate, they are investigating along with the students. They are also collaborators with each other in creating the new curricula and in helping the students undertake responsibility. Finally, they coach the students for their mastery performances before the Student Assessment Panels.

The teachers from Model High School do not serve on the panels, so they serve an advocate role rather than an evaluator role.

Because students at the Model High School demonstrate mastery of learning through performances instead of through traditional grading procedures, the school established an Administrative/Liaison Counselor position to help coordinate students' courses and transcripts with the other district high schools and with colleges.

Implications of Thematic School Restructuring

Of the 62 restructuring schools which fit the criteria of systemic restructuring, these are five examples whose restructuring is based on one or more major **integrating themes**. In all 62 schools, the structural changes are systemic (pervasive and interrelated), but the changes in the schools with **integrating themes** appear deeper and more interrelated. The schools with integrating themes also presented a more complete picture in the information they provided, compared to the less theme-based schools.

This finding suggests that a systemic restructuring effort will be easier and more effective if the new design is based on an underlying, integrating theme or two. From early indications of these schools' efforts, we hypothesize that building a restructuring process on an appropriate theme or vision may facilitate the planning and implementation of systemic change, and consequently may contribute to the creation of a more lasting, fundamentally transformed system.

General Trends in Restructuring

A second finding from the study is that a few new structural features are emerging as central to restructuring efforts across the country. We have not found universal features of restructuring; rather some general emphases of restructuring seem to be surfacing. Hence, we use the term "central features."

In presenting these central features, however, it must be cautioned that not all schools are emphasizing any of them in their restructuring efforts. In fact, some of the schools that focus heavily on an overall theme appear not to focus on one of these features because their overall theme is more idiosyncratic. Other schools may be using one or more of these central features, but only on a superficial basis. They might be implementing a feature as one of their changes, but not as one of the their central restructuring emphases.

The central features are restructuring *emphases* (hence the term central) occurring fairly frequently in our nationwide sample. They are not based on a tally of *all* the schools that are implementing that feature--for example, all the schools using team teaching. They are based instead on a count of the schools whose primary restructuring focus appears to be on that feature--e.g. team teaching--based on the information we received.

Teacher collaboration. The most frequent central feature that emerged from the study is teacher collaboration. In the information sent by schools, many schools emphasized teacher collaboration in general. In addition, two large subcategories emerged within the emphasis of teacher collaboration: site-based management and team teaching. In other words, the schools in the study which reported an emphasis on teacher collaboration appear to be focusing on teacher collaboration in general, on teacher collaboration in site-based management, or on teacher collaboration through team teaching.

Site-based management. The schools in the study emphasizing teacher collaboration through site-base management used such terms as "shared decision making," "participatory management," and "shared leadership" in describing their site-based management practices.

Although some schools are restructuring this type of collaboration to teachers, most are emphasizing collaboration among all groups, include the community, parents, and administrators. And some include students and staff. Many mechanisms and structures for making decisions and for determining various governance responsibilities were reported by the schools. One unique example was presented by Sweeney Elementary School in Santa Fe, New Mexico. A team comprised of a facilitator, three co-coordinators, a secretary, and two parents replaces the principal. The first four positions are teachers, with the facilitator a teacher who is on leave of absence and is responsible for the school's day-to-day management. For their peer evaluations, each teacher is observed by two other teachers and the facilitator.

Team teaching. Many types of team teaching were reported by schools as the major emphasis in their school restructuring effort: interdisciplinary team teaching, schools-within-a-school, teacher teams that stay with multi-aged groups for more than one year, teaching teams for at-risk students, and teaming of regular and special teachers. These schools are emphasizing teacher collaboration through the creation of teams within a school. For example, in some schools' information it appeared that facilitating teacher collaboration was their major purpose for using interdisciplinary team teaching, rather than their major purpose being to integrate the disciplines.

Heterogeneous grouping. Heterogeneous grouping of students is another central feature of restructuring efforts among schools in the SIRIUS-A study. For some schools, particularly middle and high schools, non-ability grouping (eliminating tracking) is a major emphasis. Multi-age grouping is a common restructuring feature for elementary schools.

Continuous progress / ungraded curriculum. Students progressing at their own learning or developmental pace also seems to be a central feature in restructuring efforts. This is implemented in a variety of structures, such as no grade levels in the entire school, continuous progress within two or three grade levels, and continuous progress within one grade or classroom.

Integrating disciplines / integrated learning. This central feature includes interdisciplinary team teaching where the emphasis is on providing an integrated, holistic view of the curriculum for students. Some schools have emphasized specific integrated learning theories. Guggenheim Elementary School in Chicago emphasizes integrated learning, which integrates the arts, kinesthetic activities, social and personal learning skills, and memory enhancement processes in the curriculum.

School as a center of inquiry. Another central feature for schools that are restructuring is the school as a center of inquiry. Examples of inquiry themes reported by schools in the SIRIUS-A project are: "Developing a culture of learning and professionalism for educators," "Creating an ongoing critical dialogue about all aspects of the school," "The process of learning," "A community of learners," "A community of learners and leaders," and "Thinking."

Personal student development. The "middle school concept" seems to have become an important influence in middle schools that are restructuring. The middle school concept focuses on meeting the developmental needs of the young adolescent and restructures many aspects of the school to meet those needs: time, grouping of students (usually in schools-within-a-school with teacher teams), and the role of teachers.

A major component of "middle school concept" schools and other schools from the study which focus on personal development is the "advisor/advisee program," in which students meet with small advisory groups as a regular part of the school day, either weekly or daily. In addition to helping meet students' personal and academic needs during these periods, some schools have adopted a theme of helping students develop social and group decision-making skills in these regular group meetings.

Mastery learning. In several examples of schools with strong themes or theme combinations, mastery learning is one of the major elements of the theme, and seems to make a large contribution to the extensiveness of the schools' total changes. The mastery theme seems most often to be combined with other emphases rather than being a stand-alone emphasis, although it is a stand-alone emphasis for a few schools in the study. Thus, mastery in some form is another central feature for school restructuring.

Mastery alone appears to be the central feature for restructuring in South Tama County Community School District in Tama, Iowa, and Natchez-Adams School District in Natchez, Mississippi. South Tama, whose theme is Outcome-Based Education, gives only the grades of A or B for demonstrating mastery or I for Incomplete. At Natchez-Adams, whose emphasis is mastery by objectives, all students test for mastery before moving to the next level.

Building a democratic school community. Another central feature which seems to be emerging in restructuring efforts is one of creating a "democratic community." Examples of such themes from schools in the study are "Student participation and governance in a 'Just Community,'" "A partnership approach to governance structure between students, parents, and community," and "Partnership among children, parents, staff, and community."

A proximal feature to building a community in the school is one of utilizing the outside community as part of the school. The Saturn School of Tomorrow is one example. Another is the School Without Walls in Rochester, New York, a high school where all students spend part of their day in the community in an internship with a community mentor.

Linking of schools, homes and community agencies. An emphasis of meeting the social needs of the family also seems to be emerging as a central feature of restructuring efforts. One example is Garfield Elementary School in Olympia, Washington, which focuses its restructuring on this theme. Garfield offers family support through the community mental health agency; home visits; transportation; parent classes; and medical, food bank, clothing, housing, and employment referral services. The school devotes one day per month for staff to conference with professionals from a Project Coordinating Council, which plans and implements interventions. The staff also conferences with students and parents on that day.

Another school, John Glenn Middle School in Bedford, Massachusetts, uses a Case Study Review Team of police, social workers, and juvenile probation workers to share information on problematic student and family cases.

Common Changes within Restructuring Efforts

The SIRIUS-A study yielded a wealth of information on the very large number of changes being implemented in school restructuring efforts across the country. Figure 1 lists the major categories of changes being implemented. This list represents a quantitative count, that is, those changes that seem to be most common across the country.

Curriculum	Learning / Instruction	Management/Admin	Community Service
	Mastery learning		
	Continuous progress		
	Student involvement		
	Student projects		
	Cooperative learning		
	Peer tutoring		
Advisor/advisee programs	Advisor/advisee programs		
	Individualized instruction		
	Heterogeneous grouping		
	Differentiated Staffing	Differentiated staffing	
	Staff development	Staff development	
	Teacher roles		
	Team teaching		
	Interdisciplinary		
	Schools-within-a-school		
	Learning centers/sites		
	Student assessment		
Student choice	Student choice		
Early childhood education	Early intervention programs		
	Technology		
	Use of time		
Integrating disciplines			
		Governance structures	
		Site-based management	
		Teacher team planning	
		Business involvement	
		Community involvement	Community outreach
		Parental involvement	Parent outreach

Figure 1. Types of Changes within Restructuring Efforts.

Conclusion: Differences Between Systemic Restructuring and Past Reforms

Much money has been spent on piecemeal reforms over the past 25 years, yet the quality of education has declined significantly over that time. Society is changing in ways that make our educational system obsolete and therefore ineffective in meeting the needs of children and society (Banathy, 1991). As we evolve deeper into the information age, it seems likely that this trend will not only continue, but accelerate. Does restructuring represent a different enough approach from reforming to reverse this trend? The SIRIUS-A Project is helping to answer that question by identifying some important differences between restructuring and past reform movements.

First, true restructuring involves **systemic change** in a school, including the district-level administration. Most past reforms have been piecemeal. They changed only one or a few aspects of a school, without taking a holistic consideration of how the changes will affect other aspects of the school. Systems experts have found that the parts of a social system evolve to fit well with each other. When you try to change just one part of a system, other parts will work to change it back to what it was. This explains why most educational reforms, which have thrived when external money flowed, have disappeared when the external money stopped. And similarly, it indicates that for any fundamental change to be successful, other parts of the system must change to fit well with it and support it. Consciously taking a systemic approach and making systemic changes appears to offer a much greater chance of significant and lasting improvement of education.

Second, the schools reporting the most extensive and fundamental changes have based their structural changes on underlying **integrating themes**. It would appear that the themes help to unite the changes being implemented throughout the various parts of the school by giving more meaning and direction to those participating in the restructuring. This seems likely to contribute to more cohesive and enduring change.

Third, in most of the schools, the people involved in planning their school's changes chose their themes based on the needs and values of their **community**: students, staff, parents, teachers, and other community members.[4] This required open conversations, negotiations, and real commitments from stakeholders in the school. Few past reforms have been so widespread in involving members of a school community. Such involvement is likely to contribute to more lasting and permanent change because the participants are designing the changes themselves, based on a theme that has meaning to them, which gives them ownership in the changes.

These three characteristics unique to restructuring--systemic, thematic, and stakeholder-based--distinguish restructuring from past reform movements. Systemic means the change will be pervasive and holistic, thematic means it will have meaning, and community-based means those who are affected by it have bought into it.

With systemic, fundamental change that is meaningful and that the school community itself has worked to create, restructuring appears to hold great promise for more lasting change that can result in a quantum improvement in meeting the needs of students and society in the 21st century--a radically different, post-industrial, information age.

---

[1] S.B. Merriam, Case Study Research in Education: A Qualitative Approach (San Francisco: Jossey-Bass Publishers, 1988).

[2] Indiana Curriculum Advisory Council, Indiana Schooling for the 21st Century (Indianapolis, IN: Indiana Department of Education, 1987).

[3] M.C. Moses, The Peak Performance School: A Reason for Restructuring (Westbury, NY: Wilkerson Publishing Co., 1990).

[4] This information emerged from an analysis of the change process data, which is not reported in this article.

Title:

Evaluation of Intelligent Tutoring Systems: What's been Done and What can be Done

Authors:

Michael A. Orey
Peter J. Legree
Philip D. Gillis
Estelle Bloom

Abstract

This paper explores the realm of evaluation within the domain of Intelligent Tutoring Systems (ITS). Toward that end, Littman and Soloway (1988) describe two facets of evaluation that are examined: external and internal evaluation. Described here are a variety of procedures that they did not address and some that they discounted. One of the procedures not explored is the use of Bloom's (1984) "2-Sigma Problem". Anderson, Boyle and Reiser (1985) used this procedure to externally evaluate the Lisp Tutor. Priest and Young (1988) offer a variety of procedures that can be used for internal evaluation. These procedures are presented and POSIT (an ITS for subtraction of whole numbers) is used as an example case to illustrate the evaluation procedures.

EVALUATION OF INTELLIGENT TUTORING SYSTEMS: WHAT'S BEEN DONE AND WHAT CAN BE DONE

Introduction

While there has been a great deal of expectation and a small amount of evidence for the benefits (Anderson, Boyle, & Reiser, 1985; Raghavan & Katz, 1989; Kurland, Granville, & MacLaughlin, 1990) of Intelligent Tutoring Systems (ITS), there is at least as much equivocating evidence (Center for the Study of Evaluation, 1986; Sleeman, Kelly, Martinak, Ward, and Moore, 1988; Sleeman, et.al., 1989; Stasz, 1988). The available data on these evaluations do not yet seem to support large scale expenditures for ITS implementation, but weak research methodology may be at the root of the problem. However, researchers are beginning to address the issue of how to evaluate ITS (Littman & Soloway, 1988; Priest & Young 1988). Littman and Soloway (1988) suggest that evaluation of ITS has two facets: external and internal evaluation. External evaluation addresses the question, "What is the educational impact of an ITS on students?" (p. 209). We suggest that external evaluation may be most appropriately conducted for ITS by comparing the effectiveness of the ITS to that of human tutors and regular classroom instruction, although there are some dangers in using these groups for comparisons (Clark, 1983; Hogan, 1958; Reeves, 1986).

Internal evaluation, on the other hand, addresses the question, "What is the relationship between the architecture of an ITS and its behavior?" (Littman & Soloway, 1988, p. 209). Besides architectural concerns that are common to more traditional forms of computer-based instruction (CBI) such as interface usability, components specific to an ITS need examination (e.g., student models, expert models, pedagogical models and knowledge representation). Priest and Young (1988) present four different ways of evaluating the student model within an ITS. The student model may be the most widely evaluated aspect of an ITS. Besides Priest and Young's suggestions for evaluating the student model, we suggest some ideas for evaluating the other components. In the end, a combined evaluation that rigorously tests the ITS internally and externally may begin to reveal or put an end to the great expectation for ITS.

In the following pages we describe each of these evaluation approaches and use POSIT (Orey & Burton, 1990) as a test ITS to exemplify these evaluation processes. Before we begin the description of evaluation procedures, it will help to describe the ITS used to demonstrate the procedure. In addition, a short description of the data collection is provided.

Description of POSIT

POSIT is an ITS for the tutoring of whole-number subtraction. The design of this system (see, Orey & Burton, 1990, for a detailed description of the design) is based somewhat on Anderson's (1987) ACT* theory of learning. Within this theoretical orientation, it is assumed that the learning of a cognitive skill builds from declarative knowledge. In terms of subtraction, the algorithm to be learned is made of a set of declarative information. During the development of this skill, this knowledge becomes consolidated into the specific skill (in this case, proficiency with a subtraction algorithm develops). Declarative knowledge is provided to the learner in a text form when errors occur or when the learner asks for help. For example, this is a typical message after making the Error of Omission (see Table 1 for error types) - Decrement (#1 and #2 are variables that are bound during the execution of the program) "*You have to complete the borrow into the ones place by* *taking away one of the tens. So, the correct value for this area is #1 You typed #2 Please enter the correct value.*" Although this is a very brief description of POSIT, it should suffice for the purposes of this paper. It is also necessary to provide a short description of the participants and the method used for the collection of data for this paper.

Table 1. *The four error categories and instances of errors in each category (numbers are ranked according to number of occurrences).*

Errors of Omission
1. Increment (add 10)
12. Decrement (subtract 1)
13. Both
7. Neither

Errors of Creation
8. smaller-from-larger
10. 0-N=N
9. 0-N=0
14. N-N=N
15. N-N=0 with regrouping from
16. M-S=0, with S>M

Errors of Commission
5. Increment
4. Decrement
11. Both

Other Errors
2. Fact error
6. Typing error
3. Leading 0

Data Collection

Two separate studies have been conducted on POSIT. The first was carried out in the spring of 1989 in a small town in southwest Virginia. The second was conducted in the fall of 1990 in a lower middle class suburb of Michigan. The Virginia study was conducted for the purpose of internal evaluation and the Michigan study was conducted for the purpose of external evaluation. A description of each group of participants follows.

Twelve participants were used in the Virginia evaluation. These participants were close to being 9 years old. Only those students who were having difficulties in subtraction were used. The selection of these participants was based on their teacher's recommendation and their performance on a subtraction test developed by the DEBUGGY system and reported in Van Lehn (1982). Out of a pool of 28 students who were having difficulties in subtraction, twelve participants were selected, but only ten of these completed the study. The data was collected over a five week period of time where the participants used the system about 30 minutes a day (averaging about 6 hours of total instruction). POSIT saves a great deal of information about each interaction. These data were used to demonstrate the internal evaluation procedure.

In the Michigan evaluation the entire third grade was used. This constituted 38 possible participants. One arrived late for the study and was eliminated and a second demonstrated mastery for subtraction prior to the beginning of the study which totaled 36 participants. The average age was just over 8 years. These participants were split into three groups; human tutor, computer tutor, and regular class. This yielded 12 participants for each group. The study was conducted over a five week period and the participants spent approximately 30 minutes a day working on subtraction. These data were used to demonstrate the external evaluation procedure.

The major benefit of the Michigan study was to provide appropriate feedback to the learner. The University of Georgia, Wayne State University, Apple Computer, Inc. and Oakland Elementary School, Royal Oak, Michigan, collaborated to implement this model. Both universities agreed to co-direct this project, provide technical expertise, and monitor and evaluate the program. Apple provided ten Macintosh SE's for a lab environment as well as the necessary system and utilities software.

The principal at Oakland Elementary agreed to obtain parental permission for students to participate in the POSIT program, meet with participating teachers to facilitate scheduling of students in the study, collect and record research data on a daily basis, and facilitate that data between the various sites and the University of Georgia. An integral factor was the utilization of parent volunteers to assist in implementation of this exploratory study by monitoring the various student groups involved. Most importantly, provisions were made to assure that all students who were part of the POSIT program in any of the three groups (parent volunteer, computer group, and non-computer control group) were able to receive computer instruction in a variety of curricular areas.

External Evaluations: What's been done

Although over twenty ITSs have been developed, very few of these systems have been externally evaluated. Many of these systems were developed on AI workstations that cannot be used to economically distribute training. These systems were primarily designed to demonstrate the types of interactions and approaches that could be used to deliver training. The fact that the these ITSs were not intended for widespread use may account for the paucity of formal evaluations.

Much ITS that is now being developed will run on microcomputers because of improvements in PC based software. These systems, although expensive to develop, have the potential to become cost-effective training systems because the computer programs will run on readily available microcomputers. However, because microcomputer based ITS remains expensive to develop, it is critical that performance data be collected to determine its effectiveness.

In the past, the developmental cost of ITS technology has been justified by the claim that this technology may approach the effectiveness of one-on-one human tutorial instruction (Anderson & Reiser, 1985). In fact, at least one source (Littman & Soloway, 1988) has suggested that ITS effectiveness will eventually surpass the effectiveness of the human tutor. This should have major implications for training and education because students who are taught by human tutors perform approximately two standard deviations better on instructional exams than students taught by traditional methods (Bloom, 1984). The demonstration that ITSs lead to similar gains in training effectiveness should also justify continued support for developing practical microcomputer based ITS.

Evaluating ITS technology with well designed research studies is critical to avoiding a repetition of the history of the evaluation of conventional computer aided instruction (CAI). Initial metareviews of CAI studies (Kulik, Bangert & Williams, 1983; Bangert-Downs, Kulik & Kulik, 1985; Kulik, Kulik & Shwalb, 1986) supported optimistic estimates of the effectiveness of CAI and reported effect size estimates ranging from 0.25 to 0.42. However, a metaanalysis (Clark, 1985), which eliminated poorly designed studies from Kulik and Kulik's reviews, estimated a much smaller effect size for CAI effectiveness, 0.09. This later estimate underscores the importance of using sound evaluation procedures to estimate the effectiveness of microcomputer based ITS.

Evaluations of Limited ITSs

Evaluations that have been performed on ITSs with limited scope have not been supportive of this technology. This problem is most apparent in the evaluations of older ITSs and probably reflects the fact that the first ITS applications were designed to demonstrate the instructional capabilities of ITSs. Less attention was paid to either the scope of the system, or the practical impact of the system. Performance evaluations of these systems have generally resulted in poor evaluation results. The evaluations of PROUST, PIXIE, and WEST exemplify these problems.

PROUST

PROUST was developed at Yale University to tutor students on two optional Pascal programming language problems over a semester course. The ITS was evaluated by comparing the midterm and final grades of students who opted to use PROUST for the optional homework assignments, with students who completed the homework assignments on their own. No differences were demonstrated between the two groups, although differences were demonstrated between students who completed the optional homework assignments versus those who did not (Center for the Study of Evaluation, 1986).

The major problem with the PROUST evaluation is that the ITS only supported two problems. It is unlikely that group differences would have been demonstrated had human tutors been used for the two optional assignments instead of the ITS. It is not reasonable to expect group differences supporting ITS technology given the scope of PROUST. The lack of empirical differences should not influence expectations concerning the effectiveness of ITS technology.

Pixie

Pixie, unlike WEST and PROUST, is an ITS that is still under development and will support a broad topic area, algebra instruction. Sleeman, Kelly, Martinak, Ward, and Moore (1988, 1989) describe a series of four experiments relating to the Pixie project. The Pixie research is relevant to this paper because it demonstrates that the effectiveness of a limited tutoring system is difficult to demonstrate with either computers or with human tutors.

The Pixie experiments compared the effectiveness of three tutoring strategies: Model Based Remediation (MBR), Reteaching, and a control approach. The same basic paradigm was used across the four studies, students were randomly assigned to groups of eight and were tutored according to one of the three methods for one week during their Algebra class. Thus the groups were tutored for less than five hours. The impact of the three tutoring strategies was assessed with a knowledge test following the last session.

The two experimental conditions differed in that MBR tailors instruction to the needs of the individual student while Reteaching simply reiterates the information to the student following an error. Thus MBR is ideally suited as an ITS strategy while Reteaching corresponds to a CAI approach. Students in the control condition were informed when errors were made, but no remediation was attempted.

The first study compared the effectiveness of the three approaches when presented via computers. The remaining three studies estimated the effectiveness of the three approaches when human tutors were used. The results for the human tutors and the computers were consistent in that both MBR and reteaching led to significant improvements in performance relative to the control condition, however significant differences were never demonstrated between the MBR and Reteaching conditions.

The PIXIE research indicates that five high school class periods of exposure to two tutoring approaches cannot be used to compare the effectiveness of the two approaches. This was true for instruction delivered via human tutors as well as the ITS. If the Pixie evaluation studies had used larger sample sizes, then group differences might have been demonstrated. However, effect size estimates would probably still be small.

WEST

The WEST ITS helps students play the game, How the West Was Won, which was designed to teach operator precedence. WEST's impact on learning is limited by the effectiveness of the game for teaching operator precedence and by the exposure of students to the tutor, which is less than one hour. WEST was evaluated by comparing the performance of three small groups of subjects (5 to 7 subjects per group) on mathematics tests one day after using the program for one session (Center for the Study of Evaluation, 1986).

The evaluation did not support the ITS approach in that group differences were not demonstrated. However, it is not surprising that the system did not have a measurable effect on performance because one hour of tutoring cannot be expected to have a large impact on performance. We have not found any data suggesting that one hour of tutoring, in comparison to one hour of traditional instruction, can have a large impact on performance. Thus the WEST evaluation findings are not useful in estimating the effectiveness of ITS technology and are damaging to the ITS field because they suggest that ITSs are not effective.

Evaluations of Well Developed ITSs

In contrast to ITSs designed with limited scope, ITSs developed for practical use tend to be positively evaluated. The MACH III, Lisp ITS, and Smithtown exemplify well developed systems. All three tutors were designed for an actual training requirement and cover a large amount of course material. The MACH III supports 32 hours of Army training, the Lisp ITS covers a one semester course at Carnegie Mellon University, and Smithtown corresponds to approximately one-third of an economics course.

All three of these ITSs have been favorably evaluated by the demonstration of statistically significant group differences. Furthermore, the group differences were meaningful in that each of the tutors had a substantial effect on course performance. The major problem with these studies was that the authors were not consistent in their reporting of effect size estimates or other statistics.

MACH III

The MACH III ITS evaluation report includes mean and variance estimates for students taught with the ITS, versus those taught with traditional classroom instruction. The ITS was used for approximately 32 hours of instruction and the evaluation data indicate an effect size of approximately one standard deviation (Kurland, Granville, & MacLaughlin, 1990).

Lisp ITS

The Lisp ITS was evaluated twice (Anderson & Reiser, 1985; Anderson, Boyle & Reiser, 1985). In the first evaluation, groups of 10 students learned Lisp with either a human tutor, the Lisp ITS, or through self instruction. During the first evaluation, the Lisp ITS covered 6 of the 16 chapters covered in the course, or 38 percent of the course. The first evaluation did not demonstrate group differences on course exams, however, time data indicate that the human tutor, ITS, and traditional instruction groups required different amounts of time, 11.4, 15, and 26.5 hours, to cover the same material.

In the second Lisp ITS evaluation, two groups of ten students learned Lisp with either the Lisp ITS or traditional classroom instruction. During this evaluation, the ITS covered 9 of the 16 chapters, or 56 percent of the course. A class exam indicated that the Lisp ITS students "performed 43 percent better than" the traditional classroom instruction group.

Unfortunately variance estimates were not reported for either Lisp ITS evaluation, thus it is not possible to estimate effect size. However, the time estimates and the classroom test scores are consistent with the 1.0 effect size estimated for the MACH III. The fact that differences were only demonstrated on the knowledge test for the second Lisp evaluation is most consistent with the assertion that ITS evaluations are more likely to be successful for better developed ITSs.

Smithtown

Two Smithtown evaluations are described by Raghavan and Katz (1989). The first Smithtown evaluation demonstrated that students who worked with the tutor spent much less time, 5 hours, to learn material that required 12 hours of time for students who did not have access to the tutor. This time savings estimate is consistent with time data from the Lisp ITS evaluations.

A second evaluation study demonstrated a 15 point advantage on a course test for students who had access to Smithtown during instruction as compared to students who did not. These results are consistent with the Lisp ITS data, as well as with the 1.0 standard deviation effect size estimated for the MACH III

External Evaluation: What should be done

These studies were analyzed to identify promising external evaluation procedures and criteria. The review indicated the need for three criteria that are particularly relevant to ITS evaluation studies. First, ITS evaluations should utilize three groups of students: the ITS group, a traditional instruction group, and a human tutor group. Second, only well developed ITSs should be evaluated to estimate the impact of this technology on learning. Third, ITS evaluations should utilize larger group sizes than have been used in the past. We suggest that ITS researchers ought to adopt these criteria for future ITS evaluations.

Required Groups : ITS, Human Tutor, & Traditional Instruction

Logic dictates that a minimum of three groups of subjects should be included in an ITS evaluation: a traditional classroom control, a human tutorial, and an ITS group. Although most ITS evaluations have not included these three groups, all three are critical in order to insure meaningful results. Including all three of these groups is important for a number of reasons.

ITS proponents have consistently justified ITS development on the claim that human tutoring is a highly effective means of instruction (Anderson & Reiser, 1985; Littman & Soloway, 1988). However, not all learning objectives can be assumed to be most effectively taught in a one-on-one environment. Comparing the effectiveness of a human tutor and a traditional instruction group would assess this assumption for the ITS subject matter and could justify basing ITS instruction on models of human tutoring. If human tutors cannot be demonstrated to be more effective than classroom instruction for a specific learning objective, then it is not logical to construct a system that emulates human tutors. (Obviously, it is preferable for this comparison to be made prior to developing the ITS.)

If the assumption can be made that human tutoring is more effective than traditional instruction, then including both controls in the evaluation allows the power of the research design to be assessed. Under this assumption, the experimental finding that these two groups do not differ is a Type I error and indicates problems with the experimental design. This finding could occur if the evaluation design used small sample sizes, or inaccurate performance measures.

If the human tutor and the traditional instruction groups differ, then the magnitude of the difference can be used to estimate the magnitude of the differences that can be expected between the ITS and traditional instruction groups. This information may be used to assess the adequacy of the sample sizes and performance measures used in the evaluation. This is particularly important if the evaluation does not demonstrate differences favoring the ITS group.

It is also important to compare the effectiveness of the ITS approach to the effectiveness of human tutors to allow an estimation of the extent to which the salient aspects of human tutor interactions have been encoded into the ITS. It can be expected that as ITS technologies improve, computer based tutors will approximate the effectiveness of human tutors. This comparison allows ITS developers to estimate the quality of their programs in relation to human tutor effectiveness and determine when other models are needed to further improve the effectiveness of the ITS.

Evaluating the cost-effectiveness of an ITS requires the ITS group to be compared to a traditional classroom control because in most cases, this is the least expensive instructional alternative. If the ITS cannot be shown to be either more effective than traditional instruction, or equally effective and less costly than traditional instruction, then it is not logical to develop these systems for practical applications.

Impact of the System

ITS evaluations have assessed the impact of both well developed and limited ITSs on learning. Within these evaluations, most of the well developed ITSs have been favorably evaluated, while the evaluations of limited systems have not supported ITS technology. Well developed systems are defined as systems that can be expected to have a large impact on performance, while a small impact on performance can be anticipated for a limited systems.

The magnitude of the impact of a system on performance is important for two reasons. First, demonstrating a small effect requires much greater evaluation resources than are needed to demonstrate a large effect. These resources may not always be available. Second, whether or not adequate evaluation resources are available, ITSs must be well developed in order for an evaluation to indicate a large impact on performance. This is because effect size is independent of sample size and is calculated as the ratio between the mean difference between treatment groups to the standard deviation of the observations: (Mean Difference)/(sd). This is important because ITSs that have only small effects on performance cannot be used to justify continued research and development in ITS technology. It follows that ITS evaluations should focus on well developed tutors.

POSIT was evaluated in this way in the Michigan study. The regular class (mean=358.714) was compare to the ITS (mean=296.286) and the combined standard deviation for these two groups was 79.618 with the dependent measure being the time until mastery in minutes. The resulting effect size is 0.784 (358.714-296.286/79.618). This figure is close that indicated in the MACH III evaluation. The reason for it being less than one may be due to the fact that the experiment was prematurely terminated. Because only seven of the twelve participants in the regular class instruction achieved mastery within the time scheduled for the experiment, the first seven to achieve mastery in the ITS group were used for this analysis (all twelve had achieved mastery within the scheduled time).

Group Size Requirement : Power Analyses

One problem with many evaluations is the use of small groups of subjects. Table 2 contains a listing of sample sizes for evaluations of nine ITSs. The outcome of the evaluation is also summarized in Table 2. Table 2 indicates that all of the studies used sample sizes that were less than 20. This is problematic from the standpoint of power analyses because small group sizes result in very weak ITS evaluation designs. Much larger numbers of subjects need to be included in ITS evaluation designs in order to accurately assess the impact of ITS technology.

The data summarized in Table 2 demonstrate that the issues of group size and effect size are closely linked to the probability that the system is likely to have a demonstrable effect on performance. The only evaluation studies that were favorable were those that utilized larger samples and those that have already been identified as having a substantial effect on performance (Refer to Table 2).

Table 2. *Sample sizes used in ITS evaluations.*

ITS	Sample Size Per Group	Number of Conditions	Reference[a]	Outcome[b]
Lisp ITS	10	3	1,2	1
Lisp ITS	10	2	1,2	1
PROUST	2-19	2[c]	3	0
WEST	5-7	3	3	0
MACH III	11	2	4	1
Smithtown	15[d]	2	5	1
Smithtown	15	3	5	1
Geometry Proofs	3[e]	1	6	0
Pixie	8[d]	3	7,8	0
Pixie	8[d]	3	7,8	0
Pixie	6	4	7,8	0
Pixie	8-9	3	7,8	0
ADCS	2	3	9	0
POSIT	7[f]	3	10	1
Mean	8.1	2.6		0.43

a: Refer to reference list for Table 2.
b: "1" indicates a favorable evaluation, "0" indicates no demonstrated group differences.
c: Four levels within each condition.
d: Only the total number of subjects was reported. Table values are estimates
e: Three students were pilot tested on the ITS.
f: Although the experiment began with 12 per group, it ended with only 7

References

1. Anderson & Reiser, 1985;
2. Anderson, Boyle, & Reiser, 1985
3. Center for the Study of Evaluation, 1986
4. Kurland, Granville, & MacLaughlin, 1990
5. Raghavan & Katz, 1989
6. Schofield & Evans-Rhodes, 1989
7. Sleeman, Kelly, Martinak, Ward & Moore, 1988
8. Sleeman, Kelly, Martinak, Ward & Moore, 1989
9. Crooks, Kuppin & Freedy, 1978
10. Orey, 1990

Another way to view this relationship is to consider the relationship between sample size, power, alpha probability level, and effect size. In general, as either the effect size of a treatment or the sample sizes used in a study increases, then the power of the study will also increase. Cohen (1977) has calculated power estimates for various effect sizes and sample sizes. A portion of the values reported by Cohen are listed in Table 3.

Assuming an effect size of 1.0 standard deviation unit, which is the value estimated for the MACH III effect size and is consistent with the Lisp ITS and Smithtown Effect Size, Table 3 indicates that the sample sizes used in these three successful evaluations were marginally adequate. The power associated with the studies is approximately equal to .55, .60, and .83 for the sample sizes used in the Lisp ITS, MACH III, and Smithtown evaluations. In fact, the designers of these evaluations were fortunate to have demonstrated group differences in these studies.

Table 3. *Power analysis estimates: Required sample size per group at the alpha=.05 (2-tail) level by effect size and power.*

Power	Effect Size=|MeanA-MeanB|/sigma						
	.20	.40	.60	.80	1.00	1.20	1.40
.25	84	22	10	6	5	4	3
.50	193	49	22	13	9	7	5
.60	246	62	28	16	11	8	6
.80	393	99	45	26	17	12	9
.95	651	163	73	42	27	19	14
.99	920	231	103	58	38	27	20

The impact of effect size on power is also demonstrated by comparisons within the MACH III and Lisp ITS evaluations. The MACH III evaluation compared experimental and control groups on four separate exams. Three of the exams were related to a small portion of the ITS or were psychometrically poor; group differences were not observed for these scales. On the fourth exam, the groups were significantly different at the p=.04. Thus the MACH III evaluation nearly led to a finding of no significant group differences.

The Lisp ITS was evaluated twice, the ITS covered 38 percent of the course for the first evaluation and 56 percent of the course during the second evaluation. Although the ITS groups in both evaluations spent less time solving problems, the ITS group performed significantly better than a control group only on the knowledge exam of the second evaluation. In the first evaluation, group differences were not demonstrated on a knowledge exam. It follows that given the small sample sizes used in the Lisp evaluations, the effect of the ITS on student performance could have been easily missed.

These considerations underscore the importance of evaluating well designed ITSs with large sample sizes. A conservative expectation for the purposes of power analyses is to assume an effect size of .80. Then it can be estimated that sample sizes greater than 42 per group are needed to help ensure that the ITS evaluation will not result in a Type 2 error, i.e. missing a real effect, at the 95 percent probability level (Refer to Table 3).

Other Considerations

Outcome measures

Ideally, performance data should measure both time savings, or power (percent correct). Unfortunately, these two categories conflict with each other because they are negatively correlated within individuals, i.e. the faster a person responds, the less likely that response is to be correct. Past ITS (and CAI) evaluations have used both types of measures, but the emphasis has usually been on power scales.

The strategy of attempting to demonstrate both time savings and test improvement can be problematic because the effect of the ITS is split across several dimensions. (Whereas by holding one dimension constant, e.g. time, the treatment effect would primarily affect the other dimension, i.e. percent correct.) Once the effect is split across several dimensions, demonstrating significant improvements due to the ITS on any one dimension may require a larger sample size. For this reason it is usually logical to limit variations in time on the ITS and concentrate on demonstrating differences in power scales.

Ideally, the tests should estimate the effectiveness of the ITS across several points of generalization. For example, the outcome measures should access performance on problems specifically related to the tutor, as well as on general class exams. This strategy helps to ensure that the evaluation will assess the overall effectiveness of the tutor and yield interpretable patterns of results.

Reporting Statistics

One shortcoming with many ITS evaluations is that the tests are not adequately described. Neither reliability, validity, effect size, nor variance estimates are reported. Covariates are often neither included nor discussed in the evaluation reports. Without these values it is difficult to assess the meaning of the evaluation. This is particularly true when the groups fail to be significantly different. In this case it is critical that the power of the research design be estimated.

Cognitive Science Approach

Littman and Soloway (1988) suggest that since ITS development is embedded in the field of cognitive science, a cognitive science perspective on evaluation should be taken. By a cognitive science perspective, it is implied that the system attempts to understand the cognitive processes and/or the knowledge structures (states) of the learner. Therefore, the evaluation should follow these lines. That is, the evaluation should focus on how the student is performing some process or how the student has organized some knowledge rather than on a number correct as measured on a typical achievement test. A general achievement test does not get at the level of specificity that an ITS can examine. Hence, learning is viewed as a continuous process and the system would be evaluated in terms of its ability to recognize a learner's current state and how well the learner is helped to progress from the current state to an advanced state.

However, gross analyses, as was described earlier, are also useful if they are well designed. Much of the content of the ITS can be measured on a test that was developed for the evaluation of the content area of the ITS. The processes that learners are learning from the ITS can manifest themselves both in terms of the actual knowledge and processes, but also in the product of these processes (i.e., answers on an achievement test). Therefore, there should be a distinction made between these two levels of external evaluation. At one level, the ITS is evaluated in terms of the knowledge states and cognitive processes which are a part of the design of the ITS. At the other level, the ITS is evaluated based upon performance on an achievement test specific to the domain of the ITS (as was done with POSIT in the Detroit study to determine mastery).

Table 4. *The frequency of errors in each error category on the first and last days of the treatment in Virginia.*

Errors of Omission	**First**	**Last**	**Errors of Commission**	**First**	**Last**
Increment (add 10)			Increment		1
Decrement (subtract 1)	11	1	Decrement	5	1
Both			Both		
Neither		1			
Errors of Creation			**Other Errors**		
smaller-from-larger	1		Fact error	1	
0-N=N			Typing error	5	2
0-N=0	1		Leading 0	13	1
N-N=N					
N-N=0 with regrouping from M-S=0, with S>M			Incorrect diag.	6	6

In terms of POSIT, the external evaluation at the first level would be in terms of the specific declarative knowledge (productions) that make up POSIT. Operationally, this means that there should be a reduction in the number of errors that POSIT identifies. That is, POSIT should be helping students eliminate errors in their algorithms. Table 4 shows the number of errors on the first and last days of the field test made by the ten participants that used POSIT. The first result of importance is the fact that there was an overall reduction in the number of errors (from 30 to 13 errors). However, on a cognitive science level, the results indicate that the specific error productions that the participants were using at the beginning of the field test were greatly reduced by the end of the field test. The most pronounced is the least significant -- the leading zero error type. This result is not surprising since the regular classroom teacher taught the students to fill in all spaces, so leading zeros were correct in the classroom. On the other hand, Errors of Omission were nearly as frequent as the leading zero error. This decrease would indicate that the children's procedures were being modified by their having used POSIT. Another approach might be to examine the number of steps that were carried out mentally and the time it takes between steps (this data is available, but not analyzed).

Internal Evaluation

Internal evaluation is more like research than it is like evaluation. Besides testing the functionality of the system, the theoretical assumptions about the manner in which learners learn is being tested. Internal evaluation, according to Littman and Soloway (1988), is the evaluation of the relationship between the architecture and the behavior of the ITS. The point of this evaluation is to maintain the goals of the tutor as it goes through its many revisions. A somewhat more powerful evaluation procedure was suggested by Priest and Young (1988) which might be used for internal evaluation. This procedure revolves around a cognitive science evaluation procedure in that it evaluates the internal theoretical power of the learner model.

Priest and Young (1988) suggest that there are four measures of the accuracy of the model of the learner. Those four measures are: 1) Percentage correct diagnosis; 2) Error fit measure; 3) Micro-theory evaluation quotient; and, 4) The cumulative hit curve. The first measure is the easiest, but there is no way to determine the true power of the learner model from this value. Afterall, the system may have an unique explanatory production for every error made (Young and Priest refer to this as a trivial case, which theoretically it is, but pragmatically it would be mind boggling to develop such a system). This does not contribute to a theoretical view of learning and therefore is not a powerful theory, yet the percentage correct would be 100 percent. Notice that in Figure 1 POSIT has a percentage correct of 76 percent.

The second measure, the error fit measure, is especially sensitive to false error analyses. Obviously, by subtracting wrong predictions, this value will decrease. The lower the number, the less viable the theory. Notice in Figure 1 that POSIT has a value of 0.665. The third measure specifically addresses the problem of a trivial theory. To ensure that the theory is robust enough, it is necessary to subtract the number of productions (or variant procedures) from the number of correct diagnoses. In this way, each production must be powerful enough to explain many errors or conditions. Perhaps the most powerful of the evaluation methods is the fourth: the cumulative hit curve. This curve plots the percentage of errors explained versus the number of productions used. The straight line represents the trivial case (1 to 1 correspondence between errors and productions). Figure 1 shows the cumulative hit curve for the first stage of evaluation of POSIT.

These evaluation measures help evaluate the theoretical power of a learner model within the ITS. Further, we can judge our progress by showing how the new system out performs the old system. We can use an excepted guideline. However, it is not readily apparent whether POSIT is good or bad based on these numbers. What is needed is for all systems to report these numbers so that an industry standard can be established.

Conclusions

In conclusion, no evaluation procedure is complete without some if not all of the procedures described in this paper. It is quite important to perform evaluations on an ITS both internally and externally. The external evaluation seems to be more important with regards overall instructional effectiveness while the internal evaluation is concerned with theoretical power. ITS affords the opportunity to explore theoretical issues within cognitive science as well as instructional issues within the a cognitive instructional science.

Figure 1. Priest and Young's (1988) four measures of internal theoretical power and their values for POSIT.

1) Percentage correct diagnosis.

$$\frac{\text{Number of correct diagnoses X 100}}{\text{Number of data points}} = \frac{371 \text{ X } 100}{489} = 76\% \text{ for POSIT}$$

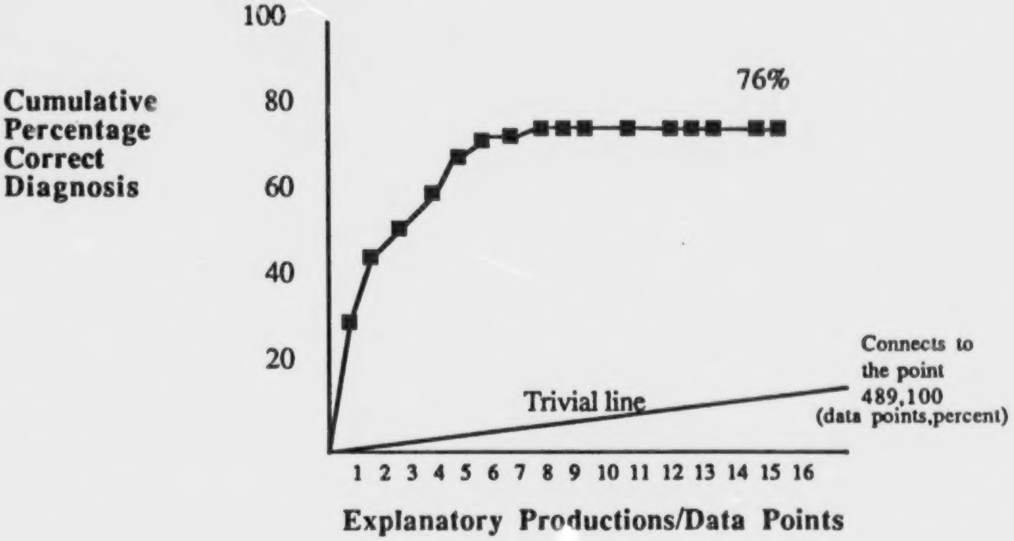
2) Error fit measure.

$$\frac{\text{Number correct diag. - Number wrong predictions}}{\text{Number of data points}} = \frac{371 - 46}{489} = 0.665 \text{ for POSIT}$$

3) Micro-theory evaluation quotient.

$$\mu = \frac{\text{Number correct diag. - Number of productions}}{\text{Number of data points}} = \frac{371 - 16}{489} = 0.726 \text{ for POSIT}$$

4) The cumulative hit curve.



References

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review. 94*(2), 192-210.

Anderson, J. R., Boyle, C. F., & Reiser, B. J. (1985). Intelligent tutoring systems. *Science. 228*, 456-462.

Anderson, J. R. & Reiser, B. J. (1985). The Lisp tutor. *Byte,10* (4), 159-175.

Bangert-Downs, R. L., Kulik, J. A. & Kulik, C. C. (1985).Effectiveness of computer based education in secondary schools. *Journal of Computer Based Instruction, 12*, 59-68.

Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4-16.

Center for the Study of Evaluation, University of California, Los Angelos. (1986). Intelligent computer aided instruction (ICAI): Formative evaluation of two systems (Army Research Institute Research Note 86-29). Alexandria, VA: U.S. Army Research Institute. (DTIC No. AD-A167 910).

Clark, R.E. (1983). Reconsidering research on learning from media. Review of *Educational Research, 53*, 445-459.

Clark, R. E. (1985). Evidence for confounding in computer-based instruction studies: Analyzing the meta-analyses. *Educational Communication Technology Journal, 33*, 249-262.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York, NY: Academic Press.

Crooks, W. H., Kuppin M. A. & Freedy, A. (1978). Application of adaptive aiding systems to computer-assisted instruction: Adaptive computer training system (ACTS) (Army Research Institute Technical Report 78-A6). Alexandria, VA: U.S. Army Research Institute. (DTIC No. AD-A056 900).

Hoban, C.F. (1958). Research on media. *AV Communications Newsletter, 6*(3), 169-178.

Kulik, J. A., Bangert, R. L. & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. *Journal of Educational Psychology, 75*, 18-26.

Kulik, C. C., Kulik, J. A. & Shwalb, B. J. (1986). The effectiveness of computer based adult education: A meta-analysis. *Journal of Educational Computing Research, 2*, 235-252.

Kurland, L. C., Granville, R. A., & MacLaughlin, D. M. (1990).*Design, development, and implementation of an intelligent tutoring system (ITS) for training radar mechanics to troubleshoot*. Unpublished manuscript, Bolt

Beranek, and Newman, Systems and Technologies Corporation, Cambridge, MA.

Littman, D., & Soloway, E. (1988). Evaluating ITSs: The cognitive science perspective. In M. C. Polson and J. J. Richardson (Eds.), *Foundations of Intelligent Tutoring Systems*, (pp.209 - 241) Hillsdale, NJ: Lawrence Erlbaum.

Orey, M. A. (1990). *Results of external evaluation.* Unpublished data.

Orey, M. A. (1989). *POSIT: Process Oriented Subtraction-Interface for Tutoring.* Unpublished dissertation, VPI & SU, Blacksburg, VA.

Orey, M. A., & Burton, J. K. (1990). POSIT: Process oriented-subtraction interface for tutoring. *Journal of Artificial Intelligence in Education, 1*(2), 77-104.

Priest, T., & Young, R. M. (1988). Methods of evaluating micro-theory systems. In John Self (Ed.), *Artificial Intelligence and Human Learning: Intelligent Computer-aided Instruction* (pp.124-137) London: Chapman and Hall Computing.

Raghavan, K. & Katz, A. (1989). Smithtown: An Intelligent Tutoring System. *Technological Horizons in Education Journal, 17* (1), 50-54.

Reeves, T.C. (1986). Research and evaluation models for the study of interactive video. *Journal of Computer-Based Instruction, 13*(4), 102-106.

Schofield J. W., Evans-Rhodes, D. & Huber, B. R. (1989).Artificial intelligence in the classroom: The impact of a computer-based tutor on teachers and students (O.N.R. Technical Report No. 3). Arlington, VA: Office of Naval Research.

Sleeman, D., Kelly, A. E., Martinak, R. Ward, R. D., & Moore, J.L. (1988). Diagnosis and remediation in the context of intelligent tutoring systems (ARI Research Note 88-66). Alexandria, VA: U.S. Army Research Institute. (DTIC No. AD-A199 024).

Sleeman, D., Kelly, A. E., Martinak, R. Ward, R. D., & Moore, J.L. (1989). Studies in the diagnosis and remediation of high school algebra students. *Cognitive Science, 13*, 551-568.

Stasz, C. (1988, April). *An intelligent tutor for basic algebra: Preliminary data on student outcomes.* Paper presented at the annual meeting of AERA in New Orleans.

VanLehn, K. (1982). Bugs are not enough: Empirical studies of bugs, impasses, and repairs in procedural skills. *Journal of Mathematical Behavior, 3*, 3-72.

Title:

Using Intelligent Tutoring Design Principles to Integrate Cognitive Theory into Computer-based Instruction

Authors:

Michael A. Orey
Wayne A. Nelson

Using Intelligent Tutoring Design Principles to Integrate Cognitive Theory into the Design of Computer-Based Instruction

The field of computer-based instruction, like other areas of instruction, has been struggling to embrace a new theoretical base. Recently, cognitive psychology has made great progress in describing the structures and processes of human cognition, and many ideas for improving instruction by relying on cognitive learning theories have been proposed (Merrill, Li & Jones, 1990a; Hannafin & Rieber, 1989a; Hannafin & Rieber, 1989b; Low, 1980; Salomon, 1985; Wildman, & Burton, 1981; Winn, 1988). Glaser (1989) has noted that cognitive theories should now address the acquisition of structures and processes, stating that "[t]he study of learning can now take its cue from this knowledge [descriptive theories], and principled investigations of instruction can be a tool of major importance for the interactive growth of learning theory and its applications" (p. 38).

The problems we are having with integration of cognitive theories into the design of computer-based instruction stem from the models we use to design instruction. Current instructional design models, since they are founded in behavioral psychology, do not adequately support the kinds of design activities and decisions necessary for instruction based in cognitive conceptions of learning. Instructional systems design models employ a systematic process for the design of instructional systems, and prescribe various activities for development of computer-based instruction such as needs assessment, development of objectives, selection of strategies, and formative evaluation. The instructional product which results from using such models typically emphasizes a frame-based approach derived from behavioral theories of human learning. For a variety of reasons, such systems and theories are now seen as largely ineffective for instruction. In order to design a new generation of computer-based instruction which incorporates principles of cognitive learning, new models for the design of computer-based instruction are necessary to guide decision-making. In addition to fitting new ideas into old models, as has been suggested by Park, Perez, & Seidel (1987), or developing a completely new model, as proposed by Merrill, Li, and Jones (1990b), we believe it is more feasible to borrow and adapt a model which has already proven effective. That model is the intelligent tutoring system.

Unlike traditional computer-based instruction, the evolution of intelligent tutoring systems better reflects the recent theoretical developments of cognitive science. Various forms of computer-based instruction today are incapable of performing well from a cognitive perspective because the design models utilized are not conducive to the development of flexible instruction which can adapt to the learner as the instruction occurs. As Winn (1988) has noted, traditional instructional design models require that all decisions about instruction be made and tested before the instruction is implemented. Winn (1989) suggests, therefore, that intelligent tutoring systems may be the instructional medium of the future because of the dynamic cognitive modeling capabilities which allow the system to make adaptive instructional decisions as the learner uses the system. In the following pages, a general model for an intelligent tutoring system is described, along with suggestions for the improvement of computer-based instruction using design principles derived from research in cognitive science.

A Model of an Intelligent Tutoring System

Intelligent tutoring systems derive historically from computer-based instruction, but since there are basic differences in theoretical perspectives, they represent two poles along a continuum of computer-based instructional systems. Differences between the two types of instructional systems are primarily structural characteristics: intelligent tutoring systems encode knowledge, while computer-based instruction encodes instructional decisions based on knowledge (Wenger, 1987). Subject matter is separated from teaching method in an intelligent tutoring system, thereby separating content from presentation techniques (Clancey, 1984). Intelligent tutoring systems are based on the idea that natural learning occurs through context-based performance, and the goal of research with intelligent tutoring systems is to develop programs which understand student misconceptions and provide appropriate instruction as a master teacher would (Loser & Kurtz, 1989).

The main advantage of the intelligent tutoring systems model, however, is that the underlying assumptions about learning and instruction differ from those of current instructional design models. Regardless of the particular model (c.f. Hartley & Sleeman, 1973; Hayes-Roth & Thorndyke, 1985; Park & Seidel, 1989; Wenger, 1987), all intelligent tutoring systems incorporate at least four major components: 1) the user interface; 2) the learner model; 3) the expert model, and; 4) pedagogic knowledge. As depicted in Figure 1, the learner and expert models are often combined into an area where diagnosis occurs through comparison between correct knowledge states (the expert model) and incorrect knowledge states (the learner model).

The *interface* provides the means for a two-way communication, where the learner is engaged in some activity while the system is interpreting the learner's activity so that a meaningful response can be made. The *expert model* is typically a database of correct knowledge states for a given domain that is organized in some form of declarative or procedural knowledge that represents the knowledge of an expert in the domain. In many intelligent tutoring systems, expert knowledge may be supplied to the system, generated dynamically, or both. The *learner model* is a representation of the errors or misconceptions that commonly occur when learners are exposed to the content, but may also include other information such as a curriculum map of the learner's sequence through the curriculum. During an instructional session, the system monitors the performance of the learner, attempting to ascertain the knowledge that the learner possesses. This diagnostic process is accomplished by comparing the learner's present knowledge state with knowledge in the expert model. The results of the comparison are then passed to the *pedagogical model*, where decisions are made about what, when, and how information is to be communicated to the learner.

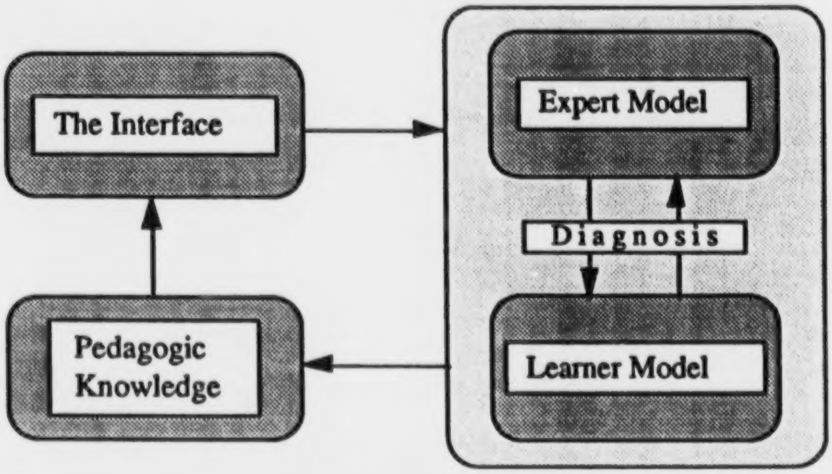


Figure 1. Components of an Intelligent Tutoring System.

Design Principles for Computer-Based Instruction

It is our contention that many of the aspects of the components of an intelligent tutoring system presented in Figure 1 can be included in any type of computer-based instructional software. Regardless of the type of software being developed, the interface should be given prime consideration, and the content needs to be clearly defined and represented (the expert model). The learner is always considered, regardless of the delivery system (the learner model) and instructional software must contain knowledge about how to provide instruction (the pedagogical model). Therefore, the design and development of computer-based instruction and intelligent tutoring systems are not vastly different. The following discussion, while not intended to be an exhaustive review of the literature, presents general guidelines for designers of computer-based instruction derived from research in cognitive science and intelligent tutoring systems.

The interface

Many principles based on cognitive theories have been proposed for user interface design. These priciples constitute the area of human-computer interaction (c.f. Card, Moran, & Newell, 1983; Norman & Draper, 1986; Shneiderman, 1987). The major problem in designing a good interface is the delay in communication with the learner that occurs between the time the system is designed and the time the learner actually uses the system. The designer must anticipate the possible actions a learner may take, and devise ways to handle the actions before the learner ever interacts with the software. This requires attention to aspects of learning, task performance, subjective satisfaction, and retention of information related to operation of the interface. Many researchers suggest that a good interface needs to be designed using some kind of iterative process for development, testing, and revision of the interface components (Shneiderman, 1987).

It is important to remember that the learner is not only learning the content, but is also learning how to operate the software. Therefore, one of the primary considerations in designing a user interface should be ease of use. Screen design, use of special function keys, menu selection, and feedback on errors are just some of the aspects which need to be considered (c. f. Jay, 1983). The key in designing these and other components is consistency. The learner should be able to rely on the same body of knowledge about operating the software as they move from one screen to another or, ideally, from one program to another. Consistent interface design will also help to reduce the cognitive load on the learner (Norman & Draper, 1986; Shneiderman, 1987), thereby freeing more processing capacity for the learning task and minimizing interference between the learning task and operation of the software (Anderson, Boyle, & Reiser, 1985; Anderson & Reiser, 1985).

Interaction style is also an important part of interface design. An interaction style which is consistent with the learners knowledge of computers is most desirable. There are many possibilities, including command-driven, menu selection, and direct manipulation interfaces. Depending on the learner's expertise, one style might be more appropriate than another. For example, if the learner is a skilled computer user, a command-driven interaction style may be the most appropriate, but if the learner is relatively inexperienced, direct manipulation of objects with a mouse may be more desirable. The selection of an interaction style should also be made based on the context of the learning activities. Menu-driven software may not be appropriate for many tasks which involve simulations, problem solving, or similar learner activities.

Metaphor can also be a powerful device in an effective interface, providing students with "comparisons which can help them learn" (Carroll & Mack, 1985; p. 40). Much of the applications software currently being marketed makes use of metaphor to help users be more productive. We have windows, trash cans, desktops, buttons, cards, folders, and pages. Such metaphors enable learners to build on their past experiences, and to map concepts in a well known domain to similar concepts and relations in a new domain. Complex metaphors may require sophisticated graphics capabilities and direct manipulation, but the benefits to the learners are worth the efforts. There are many metaphors in the classroom which are familiar to students, and which could be incorporated into an effective interface for learning.

The interface is not only a means for input and output of information, but can also supply important data about the learners. Depending on the domain, data from the interface can be used to monitor the learning process as it unfolds. If the content is process-oriented, the interface should be designed to monitor the learner's progress. For example, Orey and Burton (1990) designed an interface which supplied the system with more information about the child's subtraction process than just the answer to a problem. These data allowed the system to make decisions about the nature of the learner's subtraction errors. If the content is declarative, the interface should facilitate retention and organization. For example, the way in which an expert organizes knowledge in the field of her expertise might be depicted in some graphical form such as a content map, with the intention that the learner will develop a similar organization in obtaining domain proficiency.

The Expert Model, the Learner Model and Diagnosis

Experts in a domain organize knowledge and control problem-solving processes differently than novices (Glaser, 1989). The goal of instruction from a cognitive perspective, then, should be to replicate the knowledge structures and processes of the expert in the mind of the learner (Wildman & Burton, 1981). In order to do so, the domain expert's knowledge must be mapped into symbols a computer can store and manipulate, and presented to the learner in an organized manner. While there are many types of knowledge representation schemes, including semantic networks, production systems, scripts and/or frames, the appropriate form for knowledge representation is determined, in part, by the kinds of diagnostic procedures being implemented.

Van Lehn (1988) describes the notion of bandwidth in diagnosis, which is the correspondence between observable actions of the learner and mental states within the learner. High bandwidth is the ideal case of a one-to-one correspondence between the learner's mental states and observable actions. Low bandwidth, on the other hand, exists when there are multiple mental states between the observable actions performed by the learner. Where the system fits on the continuum between high bandwidth and low bandwidth is largely determined by the interface. For example, Orey and Burton (1990) used a process oriented interface to capture as many actions as possible while the learner performed whole number subtraction. On the other hand, Brown and Burton (1978) used a system that only examined the completed answers to an entire subtraction test. In the latter case, there are a multitude of mental states that could occur before the system is able to diagnose errors. As a result of the differences in bandwidth diagnostic levels between these two systems, the diagnostic strategies and knowledge representation were also entirely different.

The designer of computer-based instruction can make use of knowledge representation and diagnostic principles when implementing high bandwidth diagnosis. First, an interface to support this level of diagnosis needs to be designed. Second, the most common types of errors that may occur need to be identified in order to construct a partial "model" of the learner that can be used for diagnosis. The type of error made by the learner can then be determined with extended conditional statements that are checked at the point of the error (context-specific error checking). In this way, a "mini" high bandwidth diagnostic system which uses an approximation of a production system can be implemented. For example, a system for teaching hypothesis testing in statistics could present a problem, and then allow the learner to plug values from the problem statement into formulas. If the learner attempts to place a value in an incorrect position in the formula, the system could respond with appropriate instruction.

Simulation software could also benefit from monitoring and diagnosis of user errors. For example, suppose the simulation involves the operation of a nuclear power plant. The instantiation of a major error might result in the melt down of the simulated power plant, which may have dramatic effects on the learner's memory of the error. The error which caused the melt down, if properly communicated to the user, would hopefully not occur in the real world environment of running the power plant.

The Pedagogical Model

The pedagogical model contains knowledge necessary for making decisions about the teaching tactics available to the system. Since there tends to be considerable overlap between the functions of the various components of an intelligent tutoring system, the decisions and actions of the pedagogical model are highly dependent on the results of the diagnostic process. In general, the pedagogical model must decide when to present information to the learner, how to present the information, and what information to present. There have been a variety of pedagogical approaches employed in intelligent tutoring systems, but most of the current systems tend to implement only one pedagogical strategy. Intelligent tutoring systems, therefore, do not yet possess a rich repertoire of tutoring strategies from which to select. This deficiency exists, in part, because so little is known about how to teach, especially in a one-to-one setting common in tutoring (Ohlsson, 1987).

In general, pedagogical strategies are dependent on the overall context of the learning environment embodied in the intelligent tutoring system. Three general types of environments have been implemented: systems which monitor student activity within a problem-solving domain, systems which employ mixed-initiative dialogue between learner and tutor, and systems which employ guided discovery or coaching (Wenger, 1987). The choice of tutoring environment is dictated by the nature of the content to be learned, the knowledge and experience of the learner, and the assumptions about learning inherent in the underlying theory on which the system is based.

Current research with intelligent tutoring systems primarily focuses on context-based environments where students learn by working on specific, real-world problems, because cognitive theories have shown that knowledge and expertise is acquired by active application of knowledge during problem solving (Glaser, 1989). Further, learning strategies which make use of modeling of specific and appropriate procedures and strategies for problem-solving, and which strengthen existing knowledge through practice that minimizes error, have been successfully applied in intelligent tutoring systems research (Anderson, Boyle, & Reiser, 1985; White & Fredricksen, 1990). While these strategies are not significantly different from those already employed in some computer-based instruction, the use of context-based strategies supporting situated cognition (Brown, Collins, & Duguid, 1989) is becoming an important factor in the development of intelligent tutoring systems. Such strategies are not emphasized in current computer-based instruction, especially drill and practice and tutorial software.

Within a particular learning environment, a variety of teaching tactics can be implemented. Some of the possible tactics are presented in Table 1. Interested readers are referred to Wenger (1987), Ohlsson (1987), and Winne (1989) for more detailed discussions of some common teaching tactics utilized in intelligent tutoring systems. The selection and use of teaching tactics are governed by the characteristics of the learning environment, along with the general pedagogical approach. Various teaching tactics can be selected using an opportunistic approach (results of diagnosis reveal when opportunities for intervention are appropriate), or a plan-based approach (the focus of diagnosis is to monitor teaching plans and goals). Of course, a human tutor probably uses some combination of opportunistic and plan-based interventions, and research is currently focusing on the development of more sophisticated pedagogical techniques combining aspects of both pedagogical approaches (Wenger, 1987).

Table 1. *Some teaching tactics employed in intelligent tutoring systems.*

Presenting Information	Monitoring Performance	Detecting Errors
Socratic dialogue	Guided practice	Reveal errors as occur
Demonstration	Annotated practice	Marking for explanation
Priming	Hints	Probability thresholds
Associative Links	Prompt self-review	Bug repair
Curriculum Maps	Prompt self-annotation	
Issues and Examples	Evaluate hypotheses	
Answer Questions	Coaching	
	Interactive simulation (Microworld)	

Many opportunities for integrating pedagogical strategies derived from research on intelligent tutoring systems are available for designers of computer-based instruction. For example, the work on conceptual models, as discussed by Mayer (1989), can provide an effective framework for applying cognitive theories of learning in computer-based instruction systems. As an illustration of this approach, consider the study conducted by Wertheimer (1959) which examined the differences between teaching children the formulae for the area of different geometrical shapes versus teaching children the conceptual explanation of those same formulae (see Figure 2).

Initially the learner is given a paper parallelogram and a pair of scissors (Step 1 of Figure 2). The child is then instructed to cut off the triangle on one side of the parallelogram (Steps 2 and 3) and place it on the other side (Step 4). From this experience, the learner understands that the formula for the area of a parallelogram is A=B•H (where A is the area, B is the base and H is the height, or in Figure 2, A=7•3=21). Direct manipulation or animation could be used to implement this strategy in computer-based instruction (Orey, 1985). The learner could be coached through the process, learning to "grab" the triangle with a mouse and place it on the other side of the parallelogram. Such an approach allows the learner more control of the environment, and is consistent with the pedagogical strategies employed in intelligent tutoring systems.

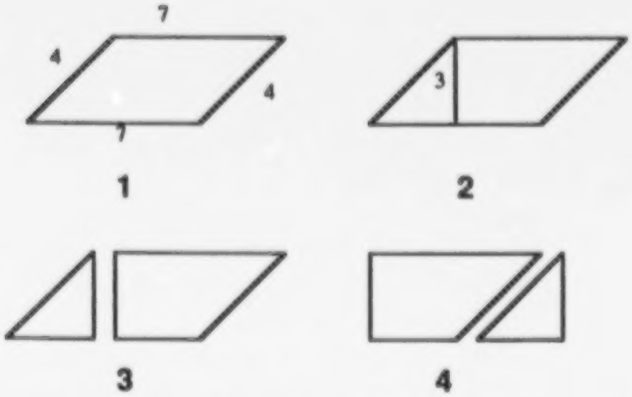


Figure 2. Wertheimer's (1959) conceptual model for teaching area.

Pedagogical knowledge, then, can be as important in computer-based instruction as it is in an intelligent tutoring system. Numerous decisions need to be made regarding teaching tactics, sequencing and media (graphics, text, sound, or video). Certainly, more traditional approaches to computer-based instruction can be used as components of the pedagogical knowledge base (see Jonassen, 1988). Metacognitive considerations may also be incorporated within the pedagogical knowledge of a computer-based instructional system. That is, communication with the learner can be used to induce reflection on the learning process or to suggest alternative problem-solving strategies (Brown, 1975). In whatever manner, it is important for the designer to remember that learners are capable of controlling their own learning to some extent (Linn & Clancey, 1990; Winne, 1989).

Conclusions

We have sketched an approach to the design of computer-based instruction derived from cognitive science, in particular, the design principles used in intelligent tutoring systems. As our analysis unfolded, we were struck by the idea that there were many similarities between intelligent tutoring systems and computer-based instruction that has been developed from a cognitive perspective. That is to say, rather than a dichotomy existing where the two poles are computer-based instruction and intelligent tutoring systems, actually there is a continuum anchored at one end by traditional computer-based instruction developed from an instructional systems design perspective (such as that found in many training settings), and at the other end by the "ideal" intelligent tutoring system. Some intelligent tutoring systems are not quite the ideal, and some computer-based instruction has been developed from a cognitive perspective. These systems overlap in the middle of the continuum.

In order for developers of computer-based instruction to integrate the findings of cognitive science into their practice, however, a new perspective is necessary. This perspective will require different assumptions about learning than those embodied in instructional design models currently used for designing computer-based instruction; assumptions that form the basis of the intelligent tutoring systems reviewed in the previous pages. It is time to apply these principles to the design and testing of computer-based instruction in real learning settings.

References

Anderson, J. R., Boyle, C. F. & Reiser, B. J. (1985). Intelligent tutoring systems. *Science, 228*, 456-462.

Anderson, J. R. & Reiser, B. J. (1985). The lisp tutor. *Byte, 10*(4), 159-175.

Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in Child Development and Behavior.* Vol 10. New York: Academic Press.

Brown, J.S., & Burton, R.R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2,* 155-192.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Erlbaum.

Carroll, J. M., & Mack, R. L. (1985). Metaphor, computing systems, and active learning. *International Journal of Man-Machine Studies, 22*, 39-57.

Clancey, W. J. (1984). Methodology for building an intelligent tutoring system. In W. Kintsch, J. R. Miller, & P. G. Polson (Eds.), *Method and tactics in cognitive science.* (pp. 51-83). Hillsdale, NJ: Erlbaum.

Hannafin, M. J., & Rieber, L. P. (1989a). Psychological foundations of instructional design for emerging computer-based instructional technologies: Part I. *Educational Technology Research and Development, 37*(2), 91-101.

Hannafin, M. J., & Rieber, L. P. (1989b). Psychological foundations of instructional design for emerging computer-based instructional technologies: Part II. *Educational Technology Research and Development, 37*(2), 102-114.

Hartley, J. R. & Sleeman, D. H. (1973). Towards intelligent teaching systems. *International Journal of Man-Machine Studies, 5*, 215-236.

Hayes-Roth, B., & Thorndyke, P. W. (1985). Paradigms for intelligent systems. *Educational Psychologist, 20*(4), 231-141.

Jay, T. B. (1983). The cognitive approach to computer courseware design and evaluation. *Educational Technology, 23*(1), 22-26.

Jonassen, D. H. (Ed.). (1988). *Instructional designs for microcomputer courseware*. Hillsdale, NJ: Erlbaum.

Linn, M. C., & Clancey, M. J. (1990). Designing instruction to take advantage of recent advances in understanding cognition. *Academic Computing, 4*(7), 20-23.

Loser, M., & Kurtz, B. (1989). Quadratic Grapher: An intelligent tutoring system for graphing quadratic equations. In H. Maurer (Ed.), *Computer Assisted Learning: Proceedings of the 2nd International Conference on Computer Assisted Learning*. (pp. 346-358). Berlin: Springer-Verlag.

Low, W. C. (1981). Changes in instructional development: The aftermath of an information processing takeover in psychology. *Journal of Instructional Development, 4*(2), 10-18.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.

Merrill, M. D., Li, Z., & Jones, M. K. (1990a). Limitations of first generation instructional design. *Educational Technology, 30*(1), 7-11.

Merrill, M. D., Li, Z., & Jones, M. K. (1990b). Second generation instructional design. *Educational Technology, 30*(2), 7-14.

Norman, D. A., & Draper, S. W. (Eds.). (1986). *User centered system design*. Hillsdale, NJ: Erlbaum.

Ohlsson, S. (1987). Some principles of intelligent tutoring. In R. W. Lawler & M. Yazdani (Eds.), *Artificial intelligence and education (Vol. 1)*. Norwood, NJ: Ablex.

Orey, M. A. (1985). *Geometric area*. Unpublished software program.

Orey, M. A., & Burton, J. K. (1990). POSIT: Process oriented subtraction-interface for tutoring. *Journal of Artificial Intelligence in Education,1*(2), 77-105.

Park, O., Perez, R. S., & Seidel, R. J. (1987). Intelligent CAI: Old wine in new bottles, or a new vintage? In G. Kearsley (Ed.), *Artificial Intelligence & Instruction: Applications and Methods*. Reading, MA: Addison Wesley.

Park, O., & Seidel, R. J. (1989). A multidisciplinary model for development of intelligent computer-assisted instruction. *Educational Technology Research and Development, 37*(3), 72-80.

Salomon, G. (1985). Information technologies: What you see is not (always) what you get. *Educational Psychologist, 20*(4), 207-216.

Shneiderman, B. (1988). *Designing the user interface: Strategies for effective human-computer interaction.* Reading, MA: Addison-Wesley.

Van Lehn, K. (1988). Student modeling. In M. Polson and J. Richardson (Eds.), *Foundations of Intelligent Tutoring Systems.* (pp. 55-78) Hillsdale, NJ: Erlbaum.

Wenger, E. (1987). *Artificial intelligence and tutoring systems.* Los Altos, CA: Morgan Kaufmann.

Wertheimer, M. (1959). *Productive thinking.* New York: Harper and Row.

White, B. Y., & Fredricksen, J. R. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence, 42*(1), 99-157.

Wildman, T. M., & Burton, J. K. (1981). Integrating learning theory with instructional design. *Journal of Instructional Development. 4*(3), 5-14.

Winn, W. (1988, January). *The theoretical foundations of educational technology and future directions for the field.* Paper presented at the annual meeting of the Association for Educational Communications and Technology. New Orleans, LA.

Winn, W. (1989). Toward a rationale and theoretical basis for educational technology. *Educational Technology Research and Development. 37*(1), 35-46.

Winne, P. H. (1989). Theories of instruction and of intelligence for designing artificially intelligent tutoring systems. *Educational Psychologist, 24*(3), 229-259.

Title:

Integrating Cognitive Theory into Gagne's Instructional Events

Authors:

Michael A. Orey
James R. Okey
Marshall G. Jones
Linda S. Stanley

Integrating Cognitive Theory into Gagné's Instructional Events

There has been a great deal of discussion about integrating cognitive theory into the development of instruction (Merrill, Li, & Jones, 1990; Orey & Nelson, 1991; Wildman & Burton, 1981; Winn, 1989). Some of the approaches have been quite radical (Merrill, Li, & Jones, 1990; Orey & Nelson, 1991; Winn, 1989). This paper will explore the possibility of integrating cognitive theory into an established development model -- Gagné's nine events of instruction. Rather than discarding the old theory and replacing it with the new -- a Popperian approach to the development of science (Popper, 1968) -- we suggest here that the field of instructional development might take the less radical form of scientific development as described by Lakatos (1978). This second form is more evolutionary than revolutionary. Hence, we will attempt to describe how current cognitive theory might evolve into our instructional development models.

Specifically, we will describe how cognitive theory might have an impact on development decisions related to Gagné's events of instruction (Gagné, Briggs, & Wager, 1988). We will use the third edition of Gagné's book to describe the most up-to-date notions of how cognitive theory impacts on this development framework. Each section of the following text examines an event of instruction. Each section includes the name of the event, the learning process associated with that instructional event, and then the textual discussion of the current cognitive theories as they relate to that event.

1. Gaining Attention. Reception of patterns of neural impulses (Gagné, Briggs, & Wager, p. 182). In terms of attention, there is a great deal of cognitive research that can have an impact on this first event of instruction. To gain and maintain attention, one needs to examine a model of attention. Norman (1976) provides the most cohesive model of attention within the information processing system. After the initial perception or pattern recognition, there is a component of memory which Norman calls "Memory Recognition" which accesses Long Term Memory (LTM). This component of memory functions on the basis of pertinence for the individual. Pertinence can be something as clear as the learner's name or as ambiguous as a scent in the air which evokes a sense of nostalgia. It should be noted in passing that this memory recognition phase is at a "deeper" level of processing than that suggested by Gagné's definition -- *Reception of patterns of neural impulses.* Gagné's definition describes perception, not attention. The question is, how does this effect the development of instruction?

Earlier attentional models focussed on the physical characteristics of the stimulus (Broadbent, 1958). Such attributes as -- the originating location of the stimulus, the volume of a sound, the brightness or different color of an image -- would cause an individual to attend to that stimulus. The later models attempted to account for such phenomena as the ability to recognize your own name being spoken in a noisy room. Norman (1976) as well as others (Shiffrin, 1976) suggested alternative models that would account for these types of phenomena.

The result is that attention is not only influenced by loud or unique stimuli, but also the context of the stimulus. The implication for this event of instruction is that attention may be achieved by using relevant stimuli as well as strange creative sounds or images. Other implications of attentional theory relate to the maintenance of attention which might be better addressed across events -- something described at the end of the paper.

2. *Informing the Learner of the Objective. Activating a process of executive control.* (Gagné, Briggs, & Wager, p. 182). Gagné implies that by providing the learner with the objectives, the learner can invoke a cognition about one's own learning. Metacognition means that one can have conscious control over one's learning or thinking (Brown, 1975). Metacognitive strategies include such things as, selecting memorial strategies for varying types of lists (choosing a mnemonic, for example), visualizing, taking notes, and selecting key words in a text to aid in comprehension. However, there is nothing inherent in a behavioral objective that would cause a learner to invoke any of these strategies. For example, a Gagné behavioral objective with its five components might be -- *[Situation:] Using a paper and pencil, [learned-capability:] list [object:] Gagné's nine events of instruction [action:] by writing them [tools and other constraints] without the use of notes, in order, and without error.* This statement establishes the goal for the student, but there is no indication about what if any executive control/metacognitive strategy should be used. If the intent of stating the objective is to *activate a process of executive control*, then it might be more appropriate to inform the learner of the task and how they might accomplish that task. This might change the above objective to -- *[Situation:] Using a paper and pencil, [learned-capability:] list [object:] Gagné's nine events of instruction [action:] by writing them [tools and other constraints] without the use of notes, in order and without error. To do this you might use a [activate a process of executive control:] mnemonic device such as an acronym or a key word method.* Afterall, not all learners are equally adept at selecting memorial strategies, so stating how the learner might learn may be quite appropriate. If the purpose of *informing the learner of the objective* is to *activate a process of executive control*, than that process ought to be stated explicitly as part of the objective. Certainly, learning and memory strategies ought to be included as part of the fourth event of instruction -- *presenting the stimulus material.*

On the other hand, Gagné also includes *expectancies* as a part of his learning model. It might be inferred that stating the objective will cause the learners to establish what behavior is *expected* of them. This definition, however, is more related to behavioral learning theory than that of cognitive theory. In cognitive psychology (we looked in several texts on cognitive psychology and the root word expectancy was not listed in the subject index, Flavell, 1985; Glover, Ronning, & Bruning, 1990; Zechmeister, & Nyberg, 1982), expectancies might be interpreted from either a prior knowledge standpoint or schema theory standpoint. In this interpretation, the learner develops expectancies about certain characteristics of their environment. The learner has come to know that when the instructor states the behavioral objectives, that the instructor is indicating a behavior that should be performed. This is different from a behavioral interpretation and several of the authors' of the current paper believe that the establishment of an expected behavior is not a bad thing for instruction. The act of goal setting is a powerful motivator in many contexts (Bandura, 1986).

3. Stimulating Recall of Prerequisite Learnings. Retrieval of prior learning to working memory: (Gagné, Briggs, & Wager, p. 182). This event is particularly compelling from the cognitive perspective. In order to develop cognitive structures (in other words, learning), one must tie what is being learned to what one already knows. Regardless, if the cognitive view is constructivist (e.g., schema theory) or information processing, the importance of prior knowledge is vital. In the schema view, the stimulation of prior knowledge will aid in the activation of appropriate schema and perhaps their modification as the result of learning. This activation of a schema aids in both the interpretation of events as well as the storage of those related memories. In the information processing view, where memory is in an associative network, the stimulation of previous knowledge identifies associations that need occur. In other words, tie the new learning to the old.

4. Presenting the Stimulus Material. Emphasizing features for selective attention. (Gagné, Briggs, & Wager, p. 182). Perhaps the most powerful application of cognitive theory to the design of instruction would be brought to bear on this event. Mayer's (1989) research on models would be effective here. Anderson's (1987) work on the development of declarative and procedural knowledge would also be a powerful construct for this event. Resnick's (1982) distinction between syntactic and semantic knowledge would effect the nature of presentation. Collins' and Loftus' (1975) network model of memory would effect how the material to be learned was organized. Metacognition (Brown, 1975) would also play an important role in the design of this event. Afterall, this is where the learning of the material takes place. The nature of the content, the organization of the content, and how the content interacts all may impact on how the information might get transferred from short term memory (STM) to LTM. Other cognitive notions that ought to impact on this event of instruction are such notions as encoding processes, knowledge as chunks of information (Miller, 1956), and parallel distributed processing (McClelland, 1988). In addition, attentional theory as well as motivational theory plays an important role in maintaining the learner's attention.

5. Providing Learner Guidance. Semantic encoding; cues for retrieval. (Gagné, Briggs, & Wager, p. 182). By semantic encoding, I assume that Gagné means elaboration, since he describes this process as one which allows the learner to move the material to LTM. It is generally accepted that information in LTM is permanent, but that people forget how to retrieve that information after awhile. Therefore, the more cues that elicit recall, the more likely the learner will be able to recall the information or the better learned the material is. There are a variety of cognitive constructs that can guide the developer when making design decisions at this phase of development. They all revolve around the notion that the more cues or links that one has for retrieving information, the more likely it is that one can do it. The first construct is that of massed versus distributed practice. A learner who is presented five opportunities to work on a specific task on five separate days (perhaps for a total of 50 minutes) will be able to perform better than a learner who is presented with those same five opportunities consecutively on the same day (also for 50 minutes). The psychological explanation for this phenomena is that the learner who receives the distributed practice may bring fresh ways of encoding the information that the massed practice learner will not do.

A second notion is that of context specific learning. If the material is presented in a variety of contexts, there will be more links to the information. Encoding specificity refers to the phenomena that occurs when the learner is provided with only a narrow context. The learner is unable to draw up that knowledge in other contexts. For example, if you learn material in one classroom, you will be able to recall the information better in that classroom than a different classroom. Ginsberg (1977) describes a child who was able to perform multi-digit subtraction at the end of the school year without error. However, at the beginning of the next school year the child performed all problems incorrectly. Upon closer examination, it was determined that the child learned to start the problem in the column closest to the piano, and the next year the piano was on the opposite side of the room.

Other factors that might be considered when developing instruction from a cognitive approach would be interference theory. Both proactive and retroactive interference can contribute to the failure of learning. In one case the new material interferes with the old and in the other the old interferes with the new. This effect is widely demonstrated with materials that are new to the learner or materials that require rote memorization. A second consideration which is directly related to interference is that of frequency. The more times you practice the better you will be given that you are provided with appropriate feedback (Event 7).

6. Eliciting Performance. Activating response organization. (Gagné, Briggs, & Wager, p. 182). It is not altogether clear what is meant by *response organization.* The nine internal processes that are described on page 11 (Gagné, Briggs, & Wager, 1988) do not include this term. However, he does depict a process of *response generation* that makes use of information from both LTM and STM. The example given is that of getting the learner to actually write the plural form of the word *appendix*. This would require the learner to retrieve from LTM the rule for performing this task and then applying the rule to this particular instance. Once the rule is applied, a response would then be generated and either voice or hand output would be initiated. This appears to be a cognitive process. However, more recent work in cognitive psychology tends to focus on process rather than product. The emphasis changes from what the observable behavior is to what that behavior indicates about the process the leaner went through to generate that particular behavior. This leads to the next event.

7. Providing Feedback about Performance Correctness. Establishing reinforcement. (Gagné, Briggs, & Wager, p. 182). Although reinforcement is a construct in learning theory, it is not a construct in cognitive learning theory. This does not mean that reinforcement suddenly does not work. It does. However, this paper is addressing cognitive theory applied to instruction and reinforcement is not a part of cognitive theory. Feedback, on the other hand, is an issue within the field and there is extensive research on the what, when and how of providing feedback. Most notably the work being done with intelligent tutoring systems (Wenger, 1987) and the research area of cognitive science (Anderson, 1987) has contributed to this area. Much of this research focuses on how the learner is responding (their solution process) and providing the specific feedback to modify the solution process so that a correct response can be generated. This does not relate to reinforcement -- a theory ladened term from operant conditioning research.

8. Assessing the Performance. Activating retrieval; making reinforcement possible. (Gagné, Briggs, & Wager, p. 182). It is not at all clear what is meant by *activating retrieval* at this point of the instruction. In terms of learning, you are trying to determine if the learner has acquired the information, organized the information appropriately, is able to retrieve the information, and/or is able to use the information in a problem solving setting. If you take Anderson's (1987) view of the acquisition of procedural knowledge, assessment must take place as close to the learning (in time) as possible. In this way, incorrect subprocedures can be corrected before they become too well established (remembered). Some other considerations, besides those implicit in the above discussion, is the the notion of recall versus recognition. Among other things, research has shown that learners who expect to be tested in a recall format (essay, problem solving test) perform better than those learners who expect a recognition format (multiple choice, matching, fill in the blank) regardless of what format the actual test takes. A second consideration might be between reconstructive and reproductive processes. With reconstructive processes, the gist of the information might be important while with reproductive processes the verbatim recall might be important.

9. Enhancing Retention and Transfer. Providing cues and strategies for retrieval. (Gagné, Briggs, & Wager, p. 182). Gagné, Briggs, and Wager (1988) suggest that retention can be enhanced by providing reviews of the content at spaced intervals (distributed practice). In order to enhance transfer, they suggest that practice be provided in widely varying contexts (again, multiple encoding links lead to more retrieval links). Transfer is widely accepted as a desirable outcome of learning. However, there are levels of transfer. For example, a math teacher might build as a rationale for doing Algebra the idea that the problem solving in the Algebra context (Two trains are heading towards each other at 60 mph and 40 mph respectively....) will transfer to other problem solving contexts (resolving marital problems). The research on this far reaching type of transfer is quite small. However, it is quite possible to help learners to transfer to similar types of problems. For example, you can do a variety of motion problems where the trains or boats or planes are heading towards or away from each other and focus the instruction on those aspects of the problem that are similar (*providing cues*).

Another aspect of transfer is the use of conceptual models (Mayer, 1989). Mayer has demonstrated several times that by using a conceptual model for the presentation of the content to naive learners, the learners are better able to remember (*retention*) and *transfer* to other areas. However, we have already suggested that this is an aspect for consideration in another event -- Presenting the stimulus material.

Not all of the events of instruction match up perfectly to cognitive theory (e.g., Brown, Collins, and Duguid's (1989) cognitive apprenticeship or microworlds from constructivistic learning theory (Harel, & Papert, 1990)), but then again not all swans are white. Instructional development is in a state of change, as it always has been, and as it probably (and hopefully) always will be. The trick is to learn from progress, and to make the change process evolutionary, and not revolutionary. We are advocating the further study of new learning theories and how and if they are complimentary to our existing theories of instructional development. By doing this we can learn how to make instruction more effective and efficient, and keep the field of instructional development in a state of growth.

References

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*(2), 192-210.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Brown, A. L. (1975). The development of memory: Knowing, knowing about knowing, and knowing how to know. In H. W. Reese (Ed.), *Advances in Child Development and Behavior*. Vol 10. New York: Academic Press.

Broadbent, D.E. (1958). *Perception and communication*. London: Pergammon.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-42.

Collins, A. M., & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Flavell, J. H. (1985). *Cognitive development* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Gagné, R. M., Briggs, L.J., & Wager, W.W. (1988). *Principles of instructional design* (3rd ed.). NY: Holt, Rinehart, and Winston.

Ginsberg, H. (1977). *Children's arithmetic*. NY: Litton.

Glover, J.A., Ronning, R.R., & Bruning, R. H. (1990). *Cognitive psychology for teachers*. NY: Macmillan.

Harel, I.,& Papert, S. (1990). Software design as a learning environment. *Interactive Learning Environments, 1*(1), 1-32.

Lakatos, I. (1978). Falsification and the methodology of scientific research programmes. In I. Lakatos and A. Musgrave (Eds.), *Criticism and the growth of knowledge.* (pp. 91-196) London: Cambridge Press.

Mayer, R. E. (1989). Models for understanding. *Review of Educational Research, 59*(1), 43-64.

McClelland, J.L. (1988). Connectionist models and psychological evidence. *Journal of Memory and Language, 27*, 107-123.

Merrill, M. D., Li, Z., & Jones, M. K. (1990). Second generation instructional design. *Educational Technology, 30*(2), 7-14.

Miller, G.A. (1956). The magical number seven, plus-or-minus two: Some limitations on our capacity for processing information. *Psychological Review, 63*, 81-97.

Norman, D.A. (1976). *Memory and attention* (2nd ed.). NY: Wiley.

Orey, M.A. & Nelson, W. A. (1990). How to apply intelligent tutoring principles to the design of computer-based instruction. In D. W. Dalton (Ed.) *Proceedings of the Association for the Development of Computer-based Instructional Systems*, (pp. 377-382), San Diego, California: ADCIS.

Popper, K. R. (1968). *Conjectures and refutations: The growth of scientific knowledge.* NY: Harper Torchbooks.

Resnick, L.B. (1982). Syntax and semantics in learning to subtract. In T.P. Carpenter, J. M. Moser, and T.A. Romberg (eds.), *Addition and subtraction: A cognitive perspective* Hillsdale, NJ: Erlbaum, 136-155.

Shiffrin, R.M. (1976). Capacity limitations in information processing, attention, and memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (pp. 64-92). Hillsdale, NJ: Erlbaum.

Wenger, E. (1987). *Artificial intelligence and tutoring systems.* Los Altos, CA: Morgan Kaufmann.

Wildman, T. M., & Burton, J. K. (1981). Integrating learning theory with instructional design. *Journal of Instructional Development.* *4*(3), 5-14.

Winn, W. (1989). Toward a rationale and theoretical basis for educational technology. *Educational Technology Research and Development.* *37*(1), 35-46.

Zechmeister, E.B., & Nyberg, S.E. (1982). *Human memory: An introduction to research and theory.* Monterey, CA: Brooks\Cole.

Title:

Adult Attitudes toward Alternative Delivery Systems and Industrial Training Outcomes

Author:

Rita C. Richey

FOUNDATIONS OF THE RESEARCH

Overview

This research is concerned with learner attitudes toward the way employee training is delivered and of the roles these convictions play in learning. It relates to media, but is not primarily concerned with media effects on the learning process, nor with the relative effectiveness of the alternative delivery systems. Rather, this research places primary emphasis on the learner—learner profile characteristics, and a range of learner attitudes—and the effects of these factors on training outcomes. While the general concepts studied here are pertinent to designing instruction for learners of all ages and for many instructional settings, this research was directed specifically towards adults participating in industrial health and safety training. It is part of a larger project designed to produce a causal model of variables which impinge upon training interventions, and as such influence adult learning.

The Adult Learner and Instructional Delivery Methodology

A traditional part of the "adult education philosophy" has been an emphasis on self-directed learning. Brookfield (1986) notes that "the development of self-directed learning capacities is perhaps the most frequently articulated aim of educators and trainers of adults (p. 40)." He continues to explain that "this self-directedness is usually defined in terms of externally observable learning activities". Knowles (1980) takes an even stronger position when he claims that the essence of adulthood is to move toward being self-directed.

Typically, this orientation towards self-directed learning is achieved through either "group-directed" or "individual-directed" delivery techniques. Examples of group-directed methodology include the use of small group discussions, role playing and simulations; individual-directed approaches would include computer-assisted instruction or programmed instruction. These are contrasted to "instructor-directed" methods, lecture being the most common example.

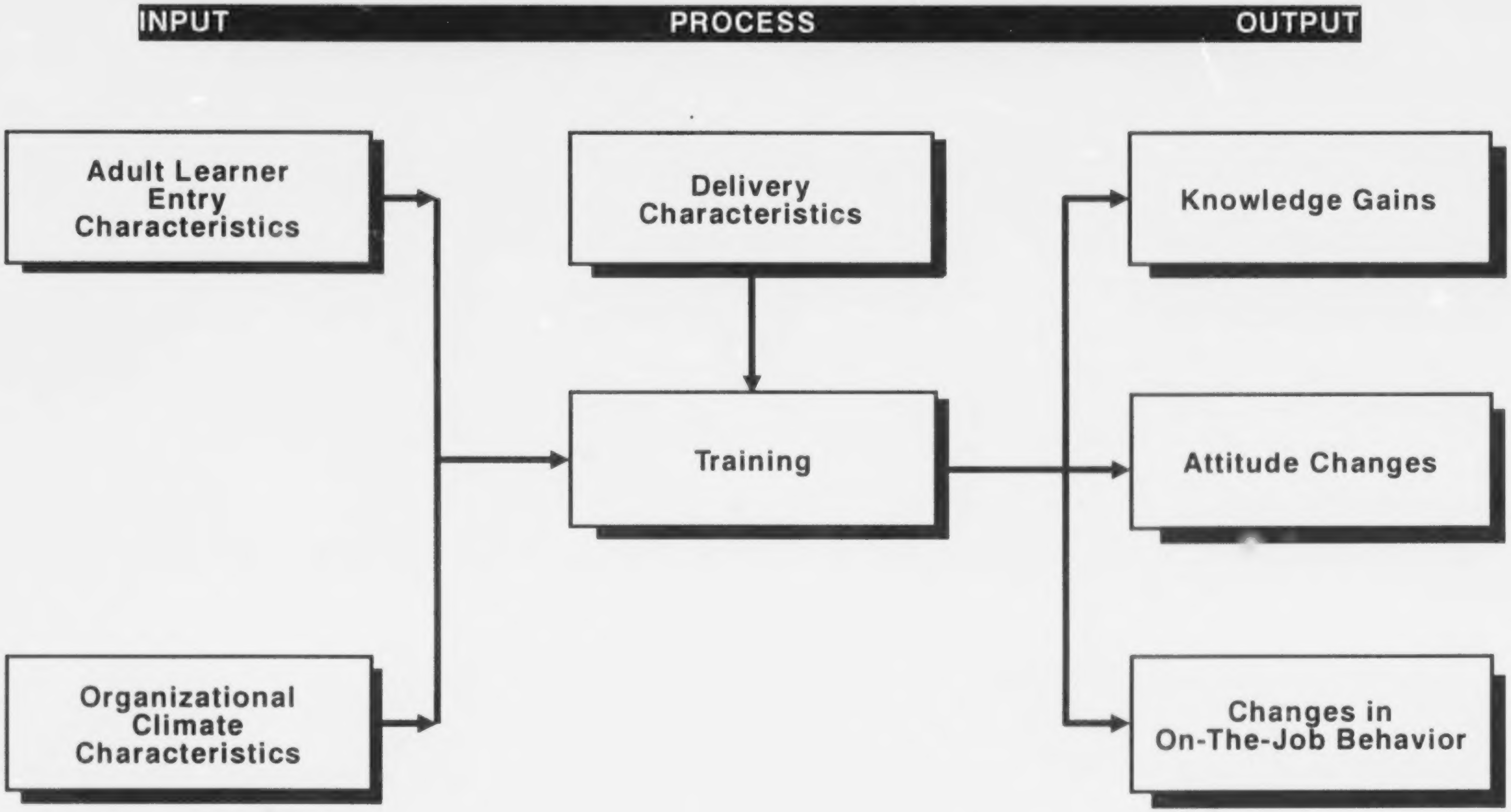
These observations are interesting in light of the common approaches to instructional delivery in corporate training. *Training* magazine's 1989 industry-wide survey showed that videotapes (used by 89.3% of responding companies) and lectures (used by 87.9%) were by far that most common vehicle for delivering instruction in the corporate training milieu. These are instructor-directed delivery systems. Group-directed methodologies were commonly used, but to a lesser extent: 58.1% used role playing; 43.9% used games and simulations in their training. In addition to these group oriented delivery systems, a smaller percentage used individualized techniques: 44.1% of the organizations with more than 100 employees engaged in computer-based training; 32.4% used non-computerized self-study programs; 11.4% used interactive video for training ("Industry Report", 1989).

Whether adult learners actually prefer those methodologies which rely on self-direction has not been empirically proven; nor is there confirmation of the assumption that self-direction is beneficial to the learning process (Caffarella and O'Donnell, 1987). In fact, the entire area of the influence of delivery system preferences on learning outcomes seems to be largely unexplored for adult learners. That is the focus of this research.

Research Framework

The general model upon which this study is based is shown in Figure 1. The model, basically an input-process-output model, shows multiple training outcomes and input from learners, environment, and delivery characteristics. In addition, it is a model which suggests causal relationships.

Figure 1
Model of Industrial Training Research


INPUT
PROCESS
OUTPUT
Adult Learner Entry Characteristics
Organizational Climate Characteristics
Delivery Characteristics
Training
Knowledge Gains
Attitude Changes
Changes in On-The-Job Behavior

The model is reminiscent of other instructional effectiveness models. However, it is not a process-product model which has been used to guide a good deal of the teaching effectiveness research. Rather than attributing learner outcomes primarily to teacher performance as do many process-product paradigms, this model provides for the possibility of multiple clusters of outcome predictors.

This model also differs from the typical instructional design procedural model which relies on a systematic orientation. This model is consistent with a more systemic orientation, an approach which emphasizes the role in the learning process of a unified whole rather than identifying and analyzing separate components as is standard in the systems approach. The creation of the "whole" with respect to instructional design and learning requires that one address a wider spectrum of variables than has typically been considered in the design of instructional programs and materials. Thus, this model is directed toward factors other than those which concern the instruction's internal structure. I have previously reported on the role of organizational climate factors in determining industrial training outcomes (Richey, 1990), and this study extends the investigation of this systemic model into an exploration of the role of learner perceptions of the delivery system.

The specific hypotheses relating to the effects of learner attitudes towards alternative delivery systems are presented in the second model shown in Figure 2. (The variables in this model are derived from the more generalized research model presented in Figure 1.) The Figure 2 model suggests there is a web of relationships between basic learner demographic characteristics and learner attitudes which partially determine the outcomes of training. Specifically, learner attitudes toward one's job, training programs in general, and the content of the training are determined by the profile characteristics and the learner's perceptions of the organizational climate. These learner attitudes, then in turn, become predictors of the learner's attitude toward the training delivery system which directly affects training outcomes.

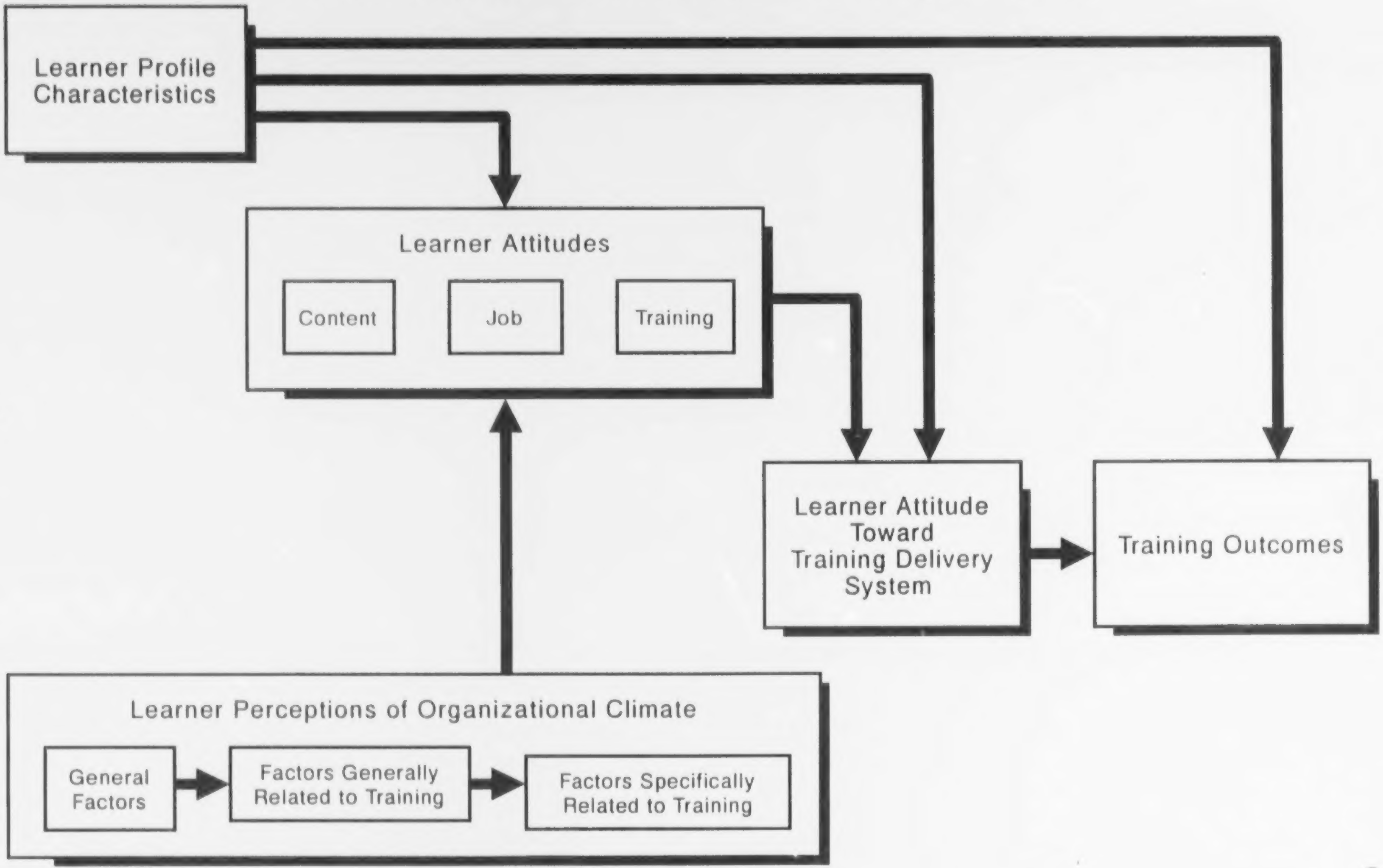
PROCEDURES

Training Framework

The two research models served as a guide for comprehensive data collection in four studies of major plant safety training programs jointly sponsored by an automotive union and a major automobile manufacturer. The first study involved a program related to energy control and power lockout (ECPL) in the plant. The topic emerged as a result of previous research and an examination of company accident records. Locking out involves shutting down the assembly line while completing diagnosis and/or repair tasks. Failure to lock out has resulted in serious injury and death. The locking out process, however, is expensive since it completely stops production. The training program on this topic was professionally designed and consisted of seven two hour sessions spanning two work weeks. Over 50,000 employees (hourly and salaried) participated. In addition, there was a one hour leadership commitment session for local plant and union management. The training was group-oriented lecture and discussion with supporting videotapes. There were pairs of trainers (one hourly and one supervisory employee), all of whom had participated in a special ECPL Train-the-Trainers program. They were not professional trainers, rather, they were released from their normal job assignments for this particular task.

The second study related to safety training operators of powered material handling vehicles (PMHV). The content focused on proper techniques for driving and controlling these vehicles in the plants. This program was offered for approximately 15,000

FIGURE 2
A Model of Hypothetical Relationships Between Delivery System Preferences and Training Outcomes


Learner Profile Characteristics
Learner Attitudes
Content
Job
Training
Learner Attitude Toward Training Delivery System
Training Outcomes
Learner Perceptions of Organizational Climate
General Factors
Factors Generally Related to Training
Factors Specifically Related to Training

plant vehicle operators and their supervisors throughout the corporation. The training consisted of four hours of group instruction supported by videotapes. Union leadership and plant management also had a one hour leadership orientation session. The program was professionally designed and delivered by trainer teams (one hourly, one supervisory) with vehicle operation experience. While they were not professional trainers, some had been trainers in the ECPL program.

Finally, an one-hour plant pedestrian safety course was studied. This program was delivered to approximately 125,000 persons, hourly and salaried, who are regular pedestrians in all plants of the company. It was a group-oriented seminar with supporting videotapes. It was professionally produced by the same firm which designed the PMHV training. Plant vehicle operators all took the pedestrian course immediately prior to the vehicle operation training. The same trainer teams were used for both PMHV pedestrian and operator training.

The final study related to the same ECPL content of the first study; however, the training format had been converted to interactive videodisc instruction. This program is required for all employees who had not participated in the previous group-oriented ECPL training, primarily new employees and persons who had returned from lay-off. The programs are being used in the plant computer labs or training facilities under the supervision of a training coordinator. Like the previous trainers, they were employees with ECPL experience, although not professional trainers. No formal computer training was provided for these trainers, although aid was available if requested. The exact same content previously incorporated into the group ECPL training is presented in the interactive video instruction. This training is in progress, and only preliminary data from this study is being presented here. The major emphasis is on the first three pieces of research.

Research Design

Not only was a common model of research variables used in these separate studies, but there were parallel data collection instruments and research designs. The research was conducted within the context of extensive evaluations of these safety training programs, and all procedures were approved by the both the union and the company.

The research employed a pre- and post-test survey design. The post-tests were administered 30-90 days after training to facilitate the collection of knowledge retention data, as well as a more realistic estimation of on-the-job behaviors.

The initial (but not the primary) thrust of the study related to describing the delivery system preferences of the adult trainees in the samples. All four trainee groups were used for this purpose. However, the major emphasis of the research concerned the effects of these learner attitudes towards the various delivery systems on training outcomes.

The dependent variables used in each parallel study were *gain scores* calculated from the pre-test and post-test measures of:

- Knowledge (based upon performance on objective tests of training content),
- Attitudes towards safety on the job (based upon self-report),
- On-the-job application of general safety precautions (based upon self-report), and
- On-the-job applications of specific behaviors taught (based upon self report).

The measures of the independent variables were essentially consistent from study to study. (In a few instances a given study had unique measures pertinent only to that particular training program. In these situations the measures can be easily categorized and compared to similar measures in the other studies.) The clusters relate to the various components of the hypothesized model in Figure 2. The specific measures are listed in Appendix A.

Population and Sample

There were two trainee populations for each study, hourly and salaried personnel. The samples were selected on a stratified basis from plants representative of the corporation primarily in terms of plant type and size. Pre-test samples were: 389 ECPL trainees, 317 operator trainees, 201 pedestrian trainees, and a preliminary sample of 34 in the ECPL/IVD training. Post-test samples were: 284 ECPL, 241 operators, and 178 pedestrians.

Five to seven plants were involved in each study, and the trainees were in randomly selected classes within each plant. (Classes are formed in the plants randomly assuring roughly equal representation from each plant department.) Table 1 describes the trainees in the two ECPL studies in terms of gender, age, educational level, and employment experience. Table 2 presents the same information for the truck operator and plant pedestrian trainees in the two remaining studies.

Data Analysis

After obtaining descriptive statistics, path analysis was used to evaluate and to estimate the dimension of the model in Figure 2. The technique enables one to estimate the causal influence of a number of variables considered simultaneously. While the hypothesized model has been formulated on the basis of theoretical expectations, path analysis permits empirical evaluation. The first three studies only have been used in the construction of the final model.

RESULTS

Summary

This research posed two major questions:

- What delivery systems do adults prefer in industrial training? (Do they really prefer self-directed instruction?)
- Does it really make any difference in training outcomes what the delivery system preferences are?

Briefly, the results are that the adult trainees in these studies consistently preferred instructor-directed delivery. Self-directed learning methodologies were the least desirable for all groups. However, there may be a tendency for the younger, more highly educated to regard these procedures in a slightly more favorable light.

In addition, attitudes towards the various delivery systems do appear to play a role in influencing a range of training outcomes. The pattern appears with respect to knowledge retention and application of general behaviors on the job; the pattern is varied in relation to facilitating attitude changes. Learners' attitudes towards delivery systems, in turn, are determined (in this research) by a complex interaction of more general attitudes toward past training, their perceptions of the organizational climate in which they work, and their own experiences.

Adult Learner Delivery System Preferences

Data were gathered on the extent to which the trainees enjoyed five different delivery systems, as well as which of those techniques were liked most and least. The methodologies considered included two group-oriented (or instructor-directed), and three individualized (or individual-directed) plans. They were:

TABLE 1
A Comparison of Characteristics of ECPL Trainees from Group Instruction and Interactive Video Instruction Programs

	Group Instruction		Interactive Video Instruction	
Characteristic	**N**	**%**	**N**	**%**
Gender				
Male	355	93.4	26	76.5
Female	25	6.6	8	23.5
Age				
25 and under	3	0.8	11	32.4
26 - 35	77	19.8	12	35.3
36 - 45	153	39.6	8	23.5
46 - 55	101	26.0	3	8.8
56 and over	50	12.9	0	0.0
Education				
Less than High School	62	15.9	3	8.8
High School	109	28.0	3	8.8
Trade School/Some College	171	44.0	8	23.5
College Degree or more	34	8.7	20	58.8
Years on Present Job				
0 - 5	176	45.2	31	91.2
6 - 10	57	14.7	1	2.9
11 - 15	48	12.0	1	2.9
16 - 20	40	10.3	0	0.0
21 and over	60	15.4	1	2.9
Years at Ford				
0 - 5	17	4.4	26	76.5
6 - 10	26	6.8	0	0.0
11 - 15	83	21.7	6	17.6
16 - 20	82	21.4	2	5.9
21 and over	175	45.7	0	0.0
Total Trainees	**389**	**100.0**	**34**	**100.0**

TABLE 2
A Comparison of PMHV Operator and Pedestrian Trainee Characteristics Using Pre-test Data

	Operator		Pedestrian	
Characteristic	N	%	N	%
Gender				
Male	310	99.4	152	77.9
Female	2	0.6	43	22.1
Age				
35 and under	19	6.3	28	14.1
36 - 46	132	43.6	78	39.4
46 - 55	119	39.3	69	34.8
56 and over	33	10.9	23	11.6
Education				
Less than High School	39	13.0	21	10.8
High School	130	43.2	93	47.9
Trade School/Some College	122	40.5	74	38.1
College Degree or More	10	3.3	6	3.1
Job Category				
Non-Production	99	32.1	15	7.7
Production	68	22.1	165	84.2
Maintenance	90	29.2	5	2.6
Supervisor	16	5.2	4	2.0
Other	35	11.4	7	3.6
Years on Present Job				
0 - 5	91	29.2	119	60.4
6 - 10	35	11.2	25	12.7
11 - 15	60	19.2	9	4.6
16 - 20	54	17.3	12	6.1
21 and over	72	23.1	32	16.2
Years at Ford				
0 - 5	6	1.9	28	14.0
6 - 10	3	1.0	4	2.0
11 - 15	32	10.3	8	4.0
16 - 20	79	25.3	43	21.5
21 and over	192	61.5	117	58.5
Total Trainees	317	100.0	201	100.0

- lecture and discussion,
- instructor led with videotape support,
- individualized instruction by a computer, (In one study differences were noted between computerized instruction with video—interactive videodisc—and computerized instruction without video support.)
- self-instructional workbooks.

Three of these delivery systems were used to some extent in the safety training related to energy control and power lockout, plant truck operator safety, and plant pedestrian safety. In the fourth study, training was delivered via interactive videodisc. There was no group-directed delivery system used in this training program.

The populations for the first three studies were comparable, except that there were essentially no females among the truck operators. The groups consisted primarily of hourly employees with a great deal of experience (nearly 20 years) with the company. There was less experience in their current positions, although it still averaged over 10 years. Most had some post-high school formal education. See Tables 1 and 2.

The preliminary data from the last study shows trainees who were younger (most were between 25 and 35), more highly educated (most have college degrees), and essentially all were new to their jobs, with the great majority new to the company. See Table 2.

Table 3 and Figure 3 show the comparisons among these groups in terms of their mean reactions to questions of the extent to which they enjoyed learning with the various delivery systems. Lecture/discussion was rated the most favorably in this measure; although learning with an instructor supported by videotape is also highly regarded. Individual-directed deliveries received the lowest ratings with the very lowest spot held by computer-based instruction, without integrated video. (The number of respondents reflects only those who indicated they had experienced the various delivery patterns.)

The preference for instructor-directed over individual-directed delivery is even more pronounced in Table 4 and Figures 4 - 7. These data were rankings—the systems liked best and least. However, in this context the system preferred by each group was the instructor with videotape. In this table it is clear that more ECPL/IVD trainees liked computerized instruction with video best than in the other three samples.

With respect to negative opinions, the first three groups clearly liked workbook self-instruction the least. For the ECPL/IVD group an equal number disliked the workbook and interactive videodisc. There are very mixed reactions to IVD instruction in this latter group. (Note: All preference data was collected before training.)

These preferences served as a key part of the model which was constructed to describe the effects of the delivery system attitudes on training outcomes.

A Model of the Effects of Delivery System Attitudes on Training Outcomes

An Overview of the Model. Figure 8 is a generalized path diagram which summarizes the replicated findings of this research. It represents conclusions derived from six validated path diagrams produced from the first three studies in this research. These paths relate to a range of training outcomes—knowledge retention, attitude change, and general application of the training content to on-the-job practices. The specific application of the skills and knowledge presented in the various training programs did not appear to be affected by trainees' delivery system preferences.

One general model, rather than one for each type of outcome, is being suggested here because the networks of causal

TABLE 3
A Comparison of Trainee Initial Attitude Toward Instructional Techniques

Instructional Technique	PMHV Operator			PMHV Pedestrian		
	N	$\bar{X}_1$	$\bar{X}_2$	N	$\bar{X}_1$	$\bar{X}_2$
Instructor-Directed						
Lecture/Discussion	257	1.914	3.086	162	1.969	3.031
Instructor/Videotape	266	1.992	3.008	154	2.058	2.942
Individual-Directed						
Computerized Instruction	194	2.650	2.350	115	2.652	2.348
Workbook Self-Instruction	250	2.684	2.316	159	2.572	2.428

$\bar{X}_1$ = Original mean
1 = Most desireable response

$\bar{X}_2$ = Converted mean
5 = Most desireable response

TABLE 3 (Continued)
A Comparison of Trainee Initial Attitude Toward Instructional Techniques

Instructional Technique	ECPL/IVD			Total Sample		
	N	$\bar{X}_1$	$\bar{X}_2$	N	$\bar{X}_1$	$\bar{X}_2$
Instructor-Directed						
Lecture/Discussion	33	1.70	3.30	452	1.886	3.114
Instructor/Videotape	30	2.00	3.00	450	2.016	2.984
Individual-Directed						
Computerized Instruction	—	—	—	309	2.654	2.346
Computer Instr/Video	23	2.39	2.61	23	2.390	2.610
Computer Instr/No Video	20	3.10	1.90	20	3.100	1.900
Workbook Self-Instruction	28	2.32	2.68	437	2.601	2.399

$\bar{X}_1$ = Original mean
1 = Most desireable response

$\bar{X}_2$ = Converted mean
5 = Most desireable response

FIGURE 3
A Graphical Comparison of Trainee Initial Attitude Toward Instructional Techniques

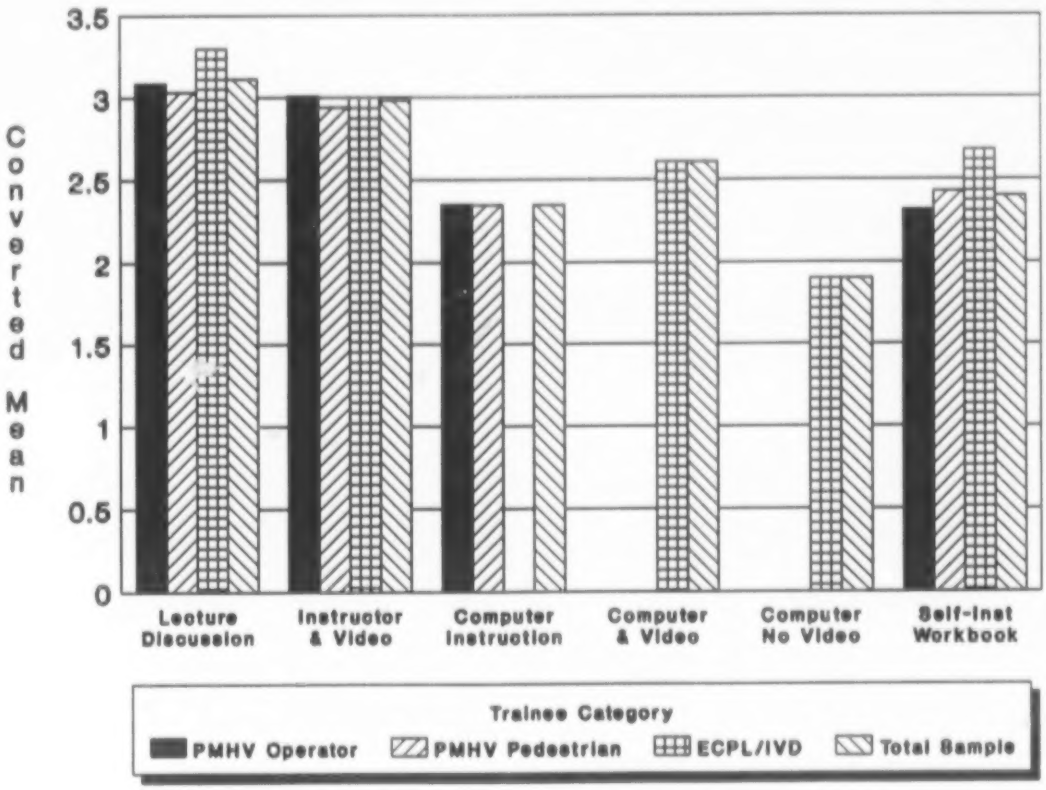


TABLE 4
A Comparison of Trainee Instructional Technique Preferences

Instructional Technique	PMHV Operator				PMHV Pedestrian			
	Liked Best		Liked Least		Liked Best		Liked Least	
	N	%	N	%	N	%	N	%
Instructor-Directed								
Lecture/Discussion	97	32.8	32	11.0	73	39.9	31	16.9
Instructor/Videotape	168	56.8	34	11.6	89	48.6	21	11.5
Individual-Directed								
Computerized Instruction	14	4.7	80	27.4	5	2.7	48	26.2
Workbook Self-Instruction	17	5.7	146	50.0	16	8.7	83	45.4
Total Sample	**296**	**100.0**	**292**	**100.0**	**183**	**100.0**	**183**	**100.0**

TABLE 4 (Continued)
A Comparison of Trainee Instructional Technique Preferences

Instructional Technique	ECPL				ECPL/IVD			
	Liked Best		Liked Least		Liked Best		Liked Least	
	N	%	N	%	N	%	N	%
Instructor-Directed								
Lecture/Discussion	59	21.3	22	7.5	5	18.5	5	14.7
Instructor/Videotape	188	67.9	23	8.3	13	48.1	10	29.4
Individual-Directed								
Computerized Instruction	24	8.7	62	22.4	—	—	—	—
Computer Instr/Video	—	—	—	—	6	22.2	9	26.5
Comp. Instr/No Video	—	—	—	—	3	11.1	1	2.9
Workbook Self-Instruction	6	2.2	188	63.7	0	0.0	9	26.5
Total Sample	**277**	**100.0**	**295**	**100.0**	**27**	**100.0**	**34**	**100.0**

FIGURE 4
A Graphical Comparison of Trainee Most Preferred Instructional Techniques

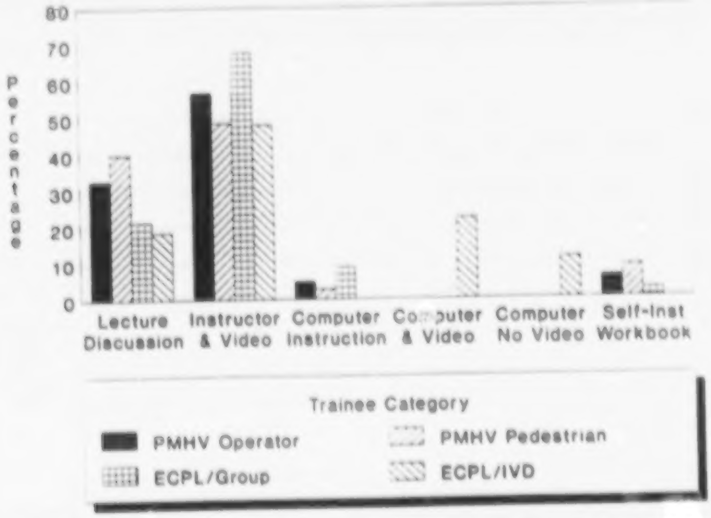


FIGURE 5
A Graphical Comparison of Trainee Least Preferred Instructional Techniques

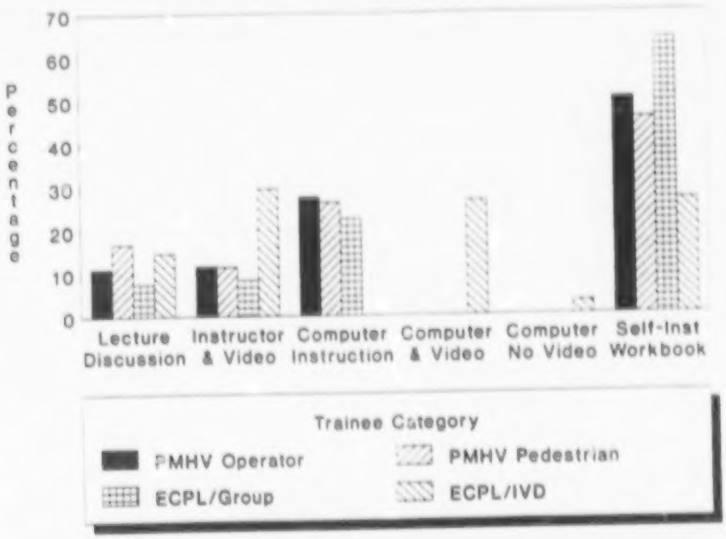


FIGURE 6
Trainee Delivery System Rankings
Most Preferred Methods (N = 783)

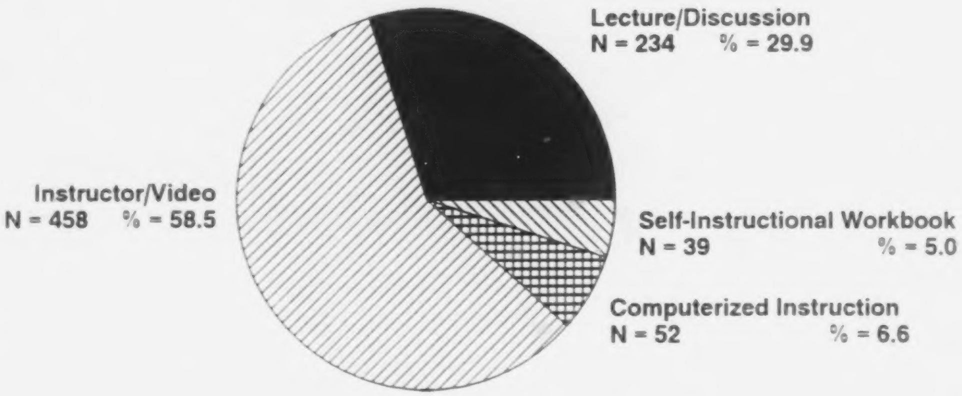


FIGURE 7
Trainee Delivery System Rankings
Least Preferred Methods (N= 804)

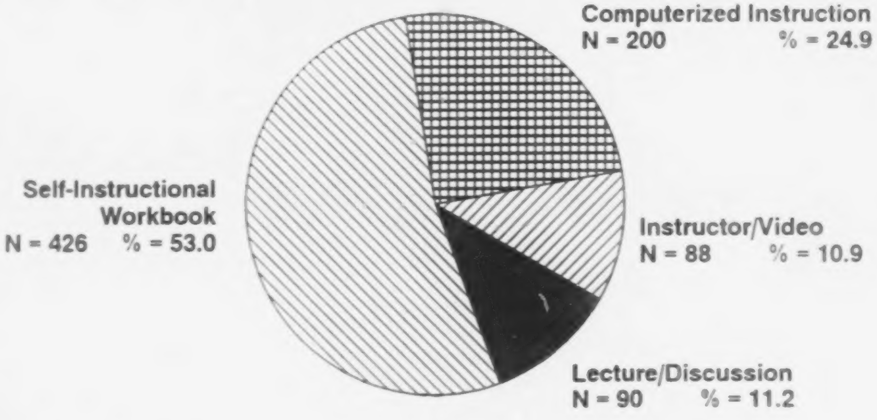
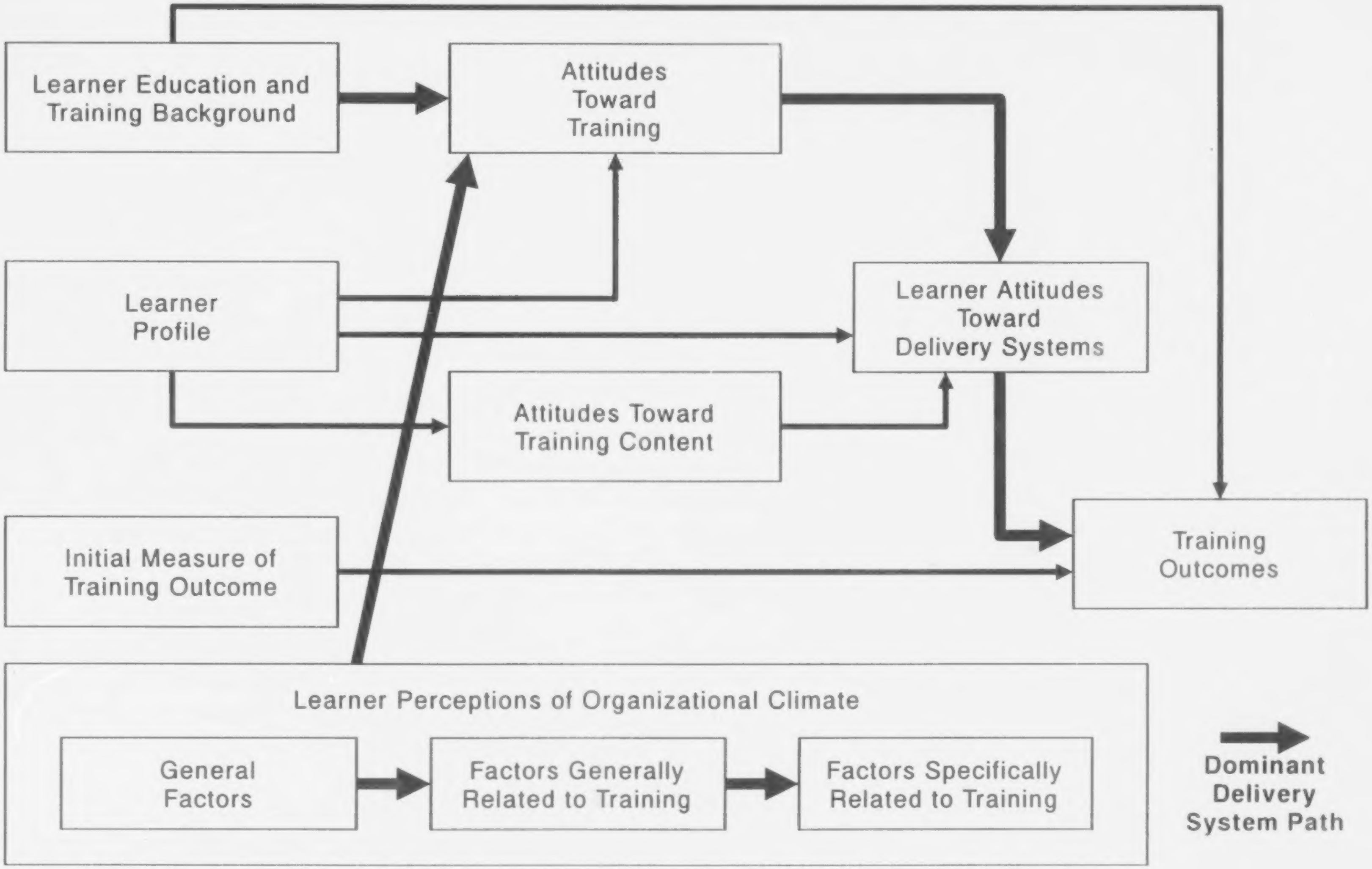


Figure 8
A General Model of Delivery System Variables Contributing to Training Outcomes


Learner Education and Training Background
Attitudes Toward Training
Learner Profile
Learner Attitudes Toward Delivery Systems
Attitudes Toward Training Content
Initial Measure of Training Outcome
Training Outcomes
Learner Perceptions of Organizational Climate
General Factors
Factors Generally Related to Training
Factors Specifically Related to Training
Dominant Delivery System Path

relationships did not seem to consistently vary in terms of training outcome. This model reflects all relationships produced by the data. However, the general model's dominant path highlights those relationships which were replicated more frequently in the separate studies.

Appendix B presents the six path diagrams upon which the general model is based. (See Figures 9 - 14.) Each of these diagrams also have dominant paths presented. Rather than representing the dominant path for that particular data set, this path shows the connections between that particular diagram and the general causal model presented in Figure 5. In addition, those measures which relate directly or indirectly to delivery system preferences have been shaded. Finally, Tables 5 - 10 present the data supporting the path diagrams in Figures 9 - 14. The measures in these diagrams match those variables listed in Appendix A.

The array of variables tested in these studies explained from 41% to 80% of the variance in the three categories of training outcomes. On the average, sixty per cent of the outcomes of these training programs can be predicted by the adult learners' entering characteristics and perceptions, most of which are directly or indirectly related to one's perceptions of the training delivery system. The factors included in the final model include:

- Initial Measure of Training Outcome,
- Perceptions of the Organizational Climate,
- Education and Training Background,
- Profile Characteristics,
- Attitudes Toward Training and the Training Content, and
- The Learner's Attitudes Toward Delivery Systems.

The Role of One's Entering Performance Level. A standard research conclusion is the best predictor of future performance is past performance. Such findings were consistent in this research. Therefore, the pre-training measure of the given outcome had the greatest influence on the gains. For example, the entering knowledge level predicted the knowledge gain; the initial safety attitude predicted the gains in safety attitudes; the behaviors demonstrated on the job prior to training predicted the changes in behavior after training.

The nature of the relationship between past and future performances, however, varies dependent upon the type of training outcome. With respect to the knowledge outcome, the relationship was expected. Those who entered training with a lower knowledge level, i.e. those with the most to gain, did demonstrate the greatest gains. For attitude and behavior changes, the pattern is reversed. Those with the highest initial measures were the more likely to demonstrate the larger gains. In other words, one apparently must have somewhat positive attitudes to begin with to be receptive to further attitude changes; and, one must already demonstrate some of the desired behaviors to be more likely to transfer training principles to the job situation. Unlike knowledge outcomes, those with the most to gain seem to be the least likely to be responsive to training. This finding was consistently replicated in the three studies.

The initial performance measure was always the major factor which determined the training outcome, accounting for the majority of the variance in the outcome measure. However, this relationship is not part of the dominant path of variables showing the relationships between adult delivery system attitudes and the major types of training outcomes. The following discussion is of those variables which do play a role in these relationships.

Learner Attitudes Toward Training Delivery. The learners' perceptions of the delivery system seem to affect more than whether they're pleased with the training; they contribute to the overall effects of the training programs. In general, if adults enjoy learning in the manner in which their training is primarily delivered, they will learn more and will be more likely to make a general transfer of the training content to their jobs. (See Figure 9 - ECPL Knowledge, Figure 10 - Pedestrian Knowledge, and Figure 13 - Pedestrian Behavior for support of this conclusion.) At times, an alternative, but supportive deduction can also be made. Disliking alternative training delivery system can also contribute to knowledge and behavior gains. (See Figure 10 - Pedestrian Knowledge and Figure 14 - Operator Behavior.) Learner perceptions of training delivery systems appear to have no effect on transferring specific skills and knowledge to the workplace.

While there are effects with respect to attitude change, a very different pattern was evident. Here disliking the primary delivery system (See Figure 12 - Operator Attitude), and liking a very different delivery system (See Figure 11 - Pedestrian Attitude) seem to contribute to gains in the desirable attitude. It may be that attitudes require an instructional methodology which facilitates more introspection than is likely with group-oriented techniques.

Attitudes Towards Training and the Trainee's Education and Training Background. There were three measures of one's attitudes towards training: Enjoying Past Training Programs, Would Volunteer for More Training, and Found Past Training Information Useful. These factors, indicative of a training motivation level, explained to a great extent the trainee's current attitudes towards the training delivery system. The relationship between attitude toward training and attitude toward delivery system is present in five of the six path diagrams. "Enjoyed past training programs" is present in each. (Only Figure 9 - ECPL Knowledge does not support this conclusion.)

The instructor/videotape format has been used frequently in the training at this corporation. Therefore, it speaks well of the training that the positive attitudes are predicted by the amount of employee training that the learners had experienced. In this context, more training leads to positive attitudes towards the experiences and, in turn, towards the delivery systems used. (See Figure 10, 11, and 13 - the three Pedestrian diagrams.) In Figure 9 - ECPL Knowledge - the amount of previous training directly influences the delivery system attitude. Level of formal education does not have the same effects consistently as amount of training. (See Figure 11 - Pedestrian Attitude and Figure 14 - Operator Behavior.)

Perceptions of Organizational Climate. The trainees' perceptions of the organizational climate directly influence their attitudes towards training. This relationship was evident in every case, except one. (See Figures 10 -14.) Three levels of organizational climate measures are included in this research: general climate measures, measures generally related to the training content (i.e. safety factors), and measures specifically related to the training content. The relationships between these levels of climate and training outcomes has previously been established (Richey, 1990); this research relates perceptions of climate to the delivery system attitudes as well. Mostly, the perceptions of general organizational climate influence attitudes towards past training; however, the other two categories share a role in determining such attitudes, as well. These relationships are evident in each supporting path diagram.

The most important factors related to employee perceptions of management actions. Eight variables influenced delivery attitudes either directly, or most frequently, indirectly. The two variables most commonly included in the various supporting path diagrams were "Supervisor's rating" and "Union/management support specific safetyrules".

The next most important categories of variables related to employee empowerment and physical working conditions, each having eighteen separate paths among the six diagrams. These two areas contained the most frequently significant variables in all of the paths. These were "Decision involvement" and "Satisfaction with physical working conditions".

A final category included fourteen paths that concerned the influence of collegial relationships. "Co-workers cooperate" and "Co-workers support safety" were the most important. Perceptions of the quality of work also influenced delivery system attitudes indirectly five times.

Learner Characteristics. Five other learner profile characteristics played an important role in explaining delivery system attitudes. Because all training programs involved health and safety issues, the prospective trainee's accident experience was assessed. The amount of accident experience seemed to promote negative attitudes towards the primary delivery system; although in one case (Figure 12 - Operator Attitude) it did relate to more enjoyment of past training.

Age and factors related to years of service were also influential. Older trainees were more likely to enjoy the instructor/videotape delivery method. (See Figure 10 - Pedestrian Knowledge and Figure 13 - Pedestrian Behavior.) Figure 10 also indicates that older workers tend to like individualized instruction by computer, which one typically would not expect. However there are conflicting signals with this diagram; it also specifies that those with more years of experience at the company tend to like computerized instruction less.

In other situations the older workers tend to have a better initial attitude toward safety. (See Figure 9 - ECPL Knowledge and Figure 11 - Pedestrian Attitude relating more years on the job to better safety attitudes.) On the other hand, the trainees with fewer years on their current jobs seemed to have a better attitude toward training in one training situation. (See Figure 14 - Operator Behavior.)

In the situation in which gender was significant, the females were more likely to enjoy learning with an instructor and videotape. (See Figures 10 and 13 - Pedestrian Knowledge and Behavior.)

Learner Attitudes Toward Training Content. The initial attitude towards safety, the general topic in each training program, influenced delivery system attitudes directly in two cases. (See Figure 9 - ECPL Knowledge and Figure 11 - Pedestrian Attitude.) However, in one instance there was a positive relationship, and it was negative in the other. No conclusive result can be determined.

DISCUSSION

This research indicates that adult attitudes towards training and the ways in which training programs are conducted do influence the fundamental success or failure of these programs, not only in terms of how much is learned, but also with respect to a generalized transfer of training principles to the workplace. A model of those factors which contribute to these training outcomes has been constructed and described. These findings have implications for

- understanding adult learning in employee training settings,
- refining the instructional design process with a more systemic orientation, and
- enhancing approaches to selecting media and learning activities.

Delivery System Preferences and Adult Learning in Employee Training

The Role of Preferences. There has been evidence that for children "student enjoyment of instructional media and their subsequent achievement were negatively correlated" (Clark and Salomon, 1986, p. 473). This appears to be a case in which adult learning patterns differ, at least in employee training situations. In the training programs studied here such enjoyment does influence achievement in both the traditional sense of knowledge acquisition, as well as in a broader definition of desired training outcomes. If these preferences are more important than being only a source of satisfaction, the question of identifying the most desirable adult delivery methodology becomes more critical.

Age and Eduction Effects. In these replicated studies, the trainees consistently prefer group instruction which is instructor-directed. This has been explained previously as an age and education-related finding. Long (1983) found that learners under 24 years of age preferred learning with peers, through direct experience, and using iconic devices. Those over 24 preferred the traditional class organization, listening and reading activities, and detailed explanations. This is not a very useful age distinction to apply to employee training, since the vast majority of trainees are over 24. The subjects of this research included 10% past the age of 55.

Knox (1977) has previously described general conclusions relating to approaches to adult learning which still have credence today. He cites the importance of one's formal preparatory educational experiences as a primary inhibitor or facilitator of adult learning. Younger adults tend to be more highly educated, and thus have had success with traditional delivery patterns. He also cites the influence of recent similar learning activities on delivery preferences.

An examination of the current findings in terms of age and education level, do not change the conclusions with respect to delivery preferences. In general, all groups preferred instructor-directed delivery. The ECPL trainees had a larger number of supervisory employees, most of whom were older, and had a higher level of formal education. An examination of the salaried employees showed little difference in their preferences when compared to hourly employees in the original set of ECPL trainees. The supervisors favored instructor with videotape support methods somewhat more than traditional lectures than did the hourly employee group.

The small preliminary set of data from the ECPL interactive-videodisc training showed employees who were very young and highly educated. They, too, preferred group instruction, but almost 25% liked IVD instruction best. This may reflect a change in current learners under 25 who have had more computer experience both in schools and the home. However, this research does not support attributing particular delivery system preferences to a given age group, or to those with a certain education level. In fact, similar preferences were identified with a population primarily of corporate managers most with engineering backgrounds (Bellinger, 1991).

Delivery System Preference and Motivation Level. Knox's (1977) emphasis on recent educational experiences may be more reflective of these results. The delivery preferences here are influenced by learners' reactions to previous training experiences, as well as the general atmosphere in their workplaces. Instructors with videotape support are used frequently in this corporation's safety training, a typical practice in large American corporations. The influence of delivery system preferences and attitudes towards past training may be an expression of the trainee's motivation level.

Typically, trainers and instructional designers motivation concerns focus on increasing the appeal of the instruction. Perhaps motivation is more closely tied to training effects. Keller (1988) has described three key approaches to incorporating motivation into instructional designs. These are person-centered, environmentally centered, and interaction-centered models. In many respects the delivery system causal model presented here represents a combination of these three approaches.

Fundamental emphasis is placed upon the personal motives and opinions of the trainee, but these attitudes are closely tied to environmental characteristics and reinforcement within the work setting. Finally, the overwhelming indication of preferences for class-based training over individually directed learning speaks to the power of a social learning context.

The missing link in this research design was a measure of preference for group-directed delivery systems. Clearly, many adult educators would feel that these attitudes could be even more positive than for the other two approaches. Such a finding would provide additional support for an interaction-centered model of motivation design.

Implications for Instructional Design

Most instructional designers are guided on a macro-level by the systems approach and on a micro-level by principles of cognitive psychology. Their work is grounded in a belief that "good design" can produce substantial learning, counterbalancing the effects of other inhibiting factors. The essence of mastery learning is faith that process, instruction and the management of instruction, can compensate for individual learner differences. The model proposed in this research says two things to the designer:

- A wide range of integrated variables affect the learning process; thus, one should consider a systemic as well as a systematic approach.
- Instructional processes may account for less than half of training outcomes.

Even as the use of instructional systems design principles is growing within the corporate training setting, others are looking ahead and becoming concerned with the effects of what they see as a blinding adherence only to procedural design models as the predominant, and often sole, guide for course and program development. Winn (1989) argues for reasoning "from basic principles of learning and instruction rather than simply following design models" (p. 36). Davies (1984) sees that the emphasis on systematic methods has resulted in a "tendency to freeze methodologies and techniques in a fixed form ... the 'one best way' approach has often reduced creative ideas to cookbook-like recipes, and promising heuristics to algorithms" (p. 9).

The model presented here shows a network of integrated variables which directly and indirectly affect training outcomes. The model does not address design process variables, such as use of feedback, effects of performance objectives, sequencing patterns, use of examples and non-examples. The model does not address delivery concerns, such as questioning techniques, reinforcement, or effects of color. While a body of literature shows the importance of factors such as these, the current research is most likely saying that those variables, even a combination of those variables, are probably not enough to explain adult learning in the workplace.

Organizational climate, one's work history, past training experiences, and delivery system attitudes account for as much knowledge retention as do formal design factors. These additional factors affect cognitive processing and are instrumental in shaping behavior on the job. While designers can say such factors are out of their control, it is difficult to ignore them when faced with the report that entry characteristics may determine up to 80% of the outcomes of training, leaving process (instruction and instructional materials) responsible for only 20% of the variance.

Delivery system preferences need to be cultivated prior to instruction in that format, especially if the methodology may be new.

This step needs to be incorporated into the instructional sequence. Its inclusion may be especially important in employee training situations with standardized delivery which allows few possibilities for trainer intervention.

Implications for Selecting Media and Delivery Methods

Modifying Current Media Selection Models. These results also have implications for selecting media and delivery systems. Reiser and Gagne (1983) present a media selection model based upon the objectives to be taught, the domain of learning to which each objective belongs, the instructional setting, and data on students' reading level. Their model systematically presents those factors typically recommended in the selection process. This research indicates that for employee training, one should also consider:

- learner attitudes towards past training and delivery systems, as well as learner abilities;
- learner related work experience and educational profiles, and
- organizational climate characteristics, as as the physical characteristics of the learning environment.

In essence, this means placing equal emphasis on learner characteristics as the model now places on content demands, media attributes, and practical considerations. For designers, this can be particularly difficult when dealing with the extreme diversity of most adult populations.

An interesting feature from one study was the fact that the older trainees liked group instruction with videotape more than lecture. The media attitudes seemed to promote both knowledge retention and behavior changes for this group. Perhaps the added properties of video creates not only a more motivating setting, but also one which highlights the information to be learned for those trainees who typically have been away from formal learning situations for some time. Younger employees, more recently in educational settings, may not need these cues.

Delivery System Selection. Romiszowski (1988) highlights the primary role of instructional methods in determining media. Delivery systems are selected first, and in turn influence media selection. A clear and replicated finding of this research relates preferred delivery with training success. However, it seems inappropriate to conclude that only familiar, well-liked delivery systems should be selected in all situations. Rather, it seems more advantageous to consider instructional interventions if new delivery systems or media are warranted.

The problems are exacerbated when considering the use of the new technologies. At times these decisions are blends of selecting both delivery system and media. Now, designers are likely to face situations with antagonistic elements: the technology is often cost effective, but the learners (and possibly the training facilitators) have had little experience with the media, or with individually directed training. Technology advantages will need to be weighed against the extra efforts needed to ensure success.

CONCLUSION

This research has produced and replicated a model showing the role attitudes toward delivery systems play in industrial training. For the most part the original hypotheses were supported. There were two areas which deviated from the original Figure 2 model. First, trainee attitudes towards their jobs do not relate to delivery system attitudes as anticipated, even though perceptions of the organizational climate play a major role. Second, delivery system preferences relate to on-the-job application of knowledge generally, but not specifically, related to the training.

656 664

The proposed causal model is complex, one which demands that designers have a thorough knowledge of their target population. It is a model which is unique in this form to employee training, rather than all adult learning environments. However, it is a model that could be useful because it was derived from real training programs with a range of typical employees. It is a model that discusses adult learning based upon empirical data, rather than only experience and hypothesis. And, as with most other research, it is a model which poses as many new questions as it answers.

REFERENCES

Bellinger, L.J. (1991). Corporate classroom instructional methods. Unpublished doctoral dissertation. Detroit, MI: Wayne State University.

Brookfield, S.D. (1986). *Understanding and facilitating adult learning*. San Francisco: Jossey-Bass Publishers.

Caffarella, R.S. and O'Donnell, J.M. (1987). Self-directed adult learning: A critical paradigm revisited. *Adult Education Quarterly*, 37, 199-211.

Clark, R.E. and Salomon G. (1986) Media in teaching. In M.C. Wittrock (Ed.) *Handbook of research on teaching* (pp. 464-478). New York: Macmillan Publishing Company.

Davies, I. (1984). Instructional development: Themata, archetypes, paradigms and models. In R.K. Bass and C.R. Dills (Ed.) *Instructional development: The state of the art*, II (pp. 8-17). Dubuque, IA: Kendall/Hunt Publishing.

Industry report, 1989. (1989, October). *Training*, pp. 31- 76.

Keller, J.M. (1988). Motivational design. In D. Unwin and R. McAleese (Ed.) *The encyclopaedia of educational media communications and technology* (pp. 406-409). New York: Greenwood Press.

Knowles, M.S. (1980). *The modern practice of adult education*. Chicago: Association Press, Follett Publishing Company.

Knox, A.B. (1977) *Adult development and learning*. San Francisco: Jossey-Bass Publishers.

Long, H.B. (1983). *Adult learning: Research and practice*. New York: Cambridge Press.

Reiser, R.A. and Gagne, R.M. (1983). *Selecting media for instruction*. Englewood Cliffs, NJ: Educational Technology Publications

Richey, R.C. (1990). The Effects of Organizational Climate Factors on Industrial Training Outcomes. In M.R. Simonson and C. Hargrave (Ed.) *Proceedings of Selected Research Paper Presentations at the 1990 Annual Convention of the Association for Educational Communication and Technology* (pp. 449-477). Ames, IA: University of Iowa.

Romiszowski, A.J. (1988). *The selection and use of instructional media*. (2nd Ed.) London: Kogan Page.

Winn, W. (1989). Toward a rationale and theoretical basis for educational technology. *Educational Technology Research and Development, 37*(1), 35-46.

APPENDIX A

MEASUREMENT OF VARIABLES

All measures listed below were used in each of the four studies unless otherwise noted. The following key relates those unique variables to the appropriate study:

* Energy Control and Power Lockout Safety Training/Group

** Energy Control and Power Lockout Safety Training/IVD

Plant Pedestrian Safety Training

Truck Operator Safety Training

Only measures included on path diagrams are listed.

EXOGENOUS VARIABLES

Learner Demographic Profile Characteristics

1. Age. In years.
2. Gender. A dummy variable, 1 = male; 2 = female.
3. Education level. 1-5 graduated levels.
4. Amount of previous training. *1-5 graduated levels; **,#, and ## 1-6 graduated levels.

Learner Work Experience Characteristics

1. Years on present job. 1-5 graduated levels.
2. Years employed by company. 1-5 graduated levels.
3. Accident Experience. *1 = yes; 2 = no. **,#, and ## 1= none; 2 = known someone in accident; 3 = minor accident; 4 = major accident.

General Organizational Climate Characteristics

Employee Perceptions of Management Actions

1.Supervisor's rating. 1-5 Very good to very poor.

Employee Empowerment

1. Involvement in decision making. 1-5 Very satisfied to very dissatisfied.
2. Encouraged to devise new work methods. 1-5 Strongly agree to strongly disagree.

Collegial Relationships

1. Co-workers cooperate. 1-5 Strongly agree to strongly disagree.

Physical Conditions

1. Physical working conditions. 1-5 Strongly agree to strongly disagree.
2. Conditions promote productivity. 1-5 Strongly agree to strongly disagree.

Quality of Work

1. Quality of work. 1-5 Strongly agree to strongly disagree.

ENDOGENOUS VARIABLES

Learner Attitudes

1. Job satisfaction.1-5 Strongly agree to strongly disagree.
2. Safety attitude. 1-5 Strongly agree to strongly disagree.

Attitude Toward Training

1. Volunteer for other training. 1-5 Strongly agree to strongly disagree. (#. ##, and **); A dummy variable 0 = no 1 = yes *
2. Information from past training useful. 1-5 Strongly agree to strongly disagree. (#. ##, and **); A dummy variable 0 = no 1 = yes *
3. Enjoy past training. 1-5 Strongly agree to strongly disagree. (#. ##, and **); A dummy variable 0 = no 1 = yes *

Organizational Climate Generally Related to Training

Employee Perceptions of Management Actions

1. Operation more important than safety. 1-5 Strongly agree to strongly disagree. 5 = most desirable (# and ##)
2. Plant manager aware of safety issues. 1-5 Strongly agree to strongly disagree.
3. Union/management support safety. 1-5 Strongly agree to strongly disagree.
4. Supervisor corrects safety hazards quickly. 1-5 Strongly agree to strongly disagree.
5. Safety top plant priority. 1-5 Strongly agree to strongly disagree.

Collegial Relationships

1. Co-workers support safety. 1-5 Strongly agree to strongly disagree.

Organizational Climate Specifically Related to Training

Employee Perceptions of Management Actions

1. Union/management support lockout/truck operator/pedestrian safety rules. 1-5 Strongly agree to strongly disagree.
2. Accountable for not following lockout/truck operator/pedestrian safety rules. 1-5 Strongly agree to strongly disagree.
3. Lockout not a common practice.1-5 Strongly agree to strongly disagree. (* and **)

Collegial Relationships

1. Pedestrians/Truck operators follow safety rules. 1-5 Strongly agree to strongly disagree. (# and ##)

Physical Conditions

1. Shortage of locks.1-5 Strongly agree to strongly disagree. (* and **)
2. Impossible to lockout on my job. 1-5 Strongly agree to strongly disagree. (* and **)

Learner Attitudes Toward Delivery System

1. Like learning with instructor/videotape best. A dummy variable. 0=no 1=yes
2. Enjoy learning with instructor/videotape. 0=Never experienced 1-5 Strongly agree to strongly disagree. (#,##, and **)
3. Enjoy learning with lecture/discussion. 0 = Never experienced 1-5 Strongly agree to strongly disagree. (#,##, and **)
4. Enjoy learning with self-instructional workbook. 0=Never experienced 1-5 Strongly agree to strongly disagree. (#,##, and **)
5. Enjoy learning with individualized instruction with computer. 0 = Never experienced 1-5 Strongly agree to strongly disagree. (#, and ##)

DEPENDENT VARIABLES

1. Gains in Knowledge Retention. The pre-test/post-test differences between the sums of the correct answers of those objective test items covering training content. (ECPL group and interactive video, 23 items; Truck Operator, 25 items; Plant Pedestrian, 14 items)

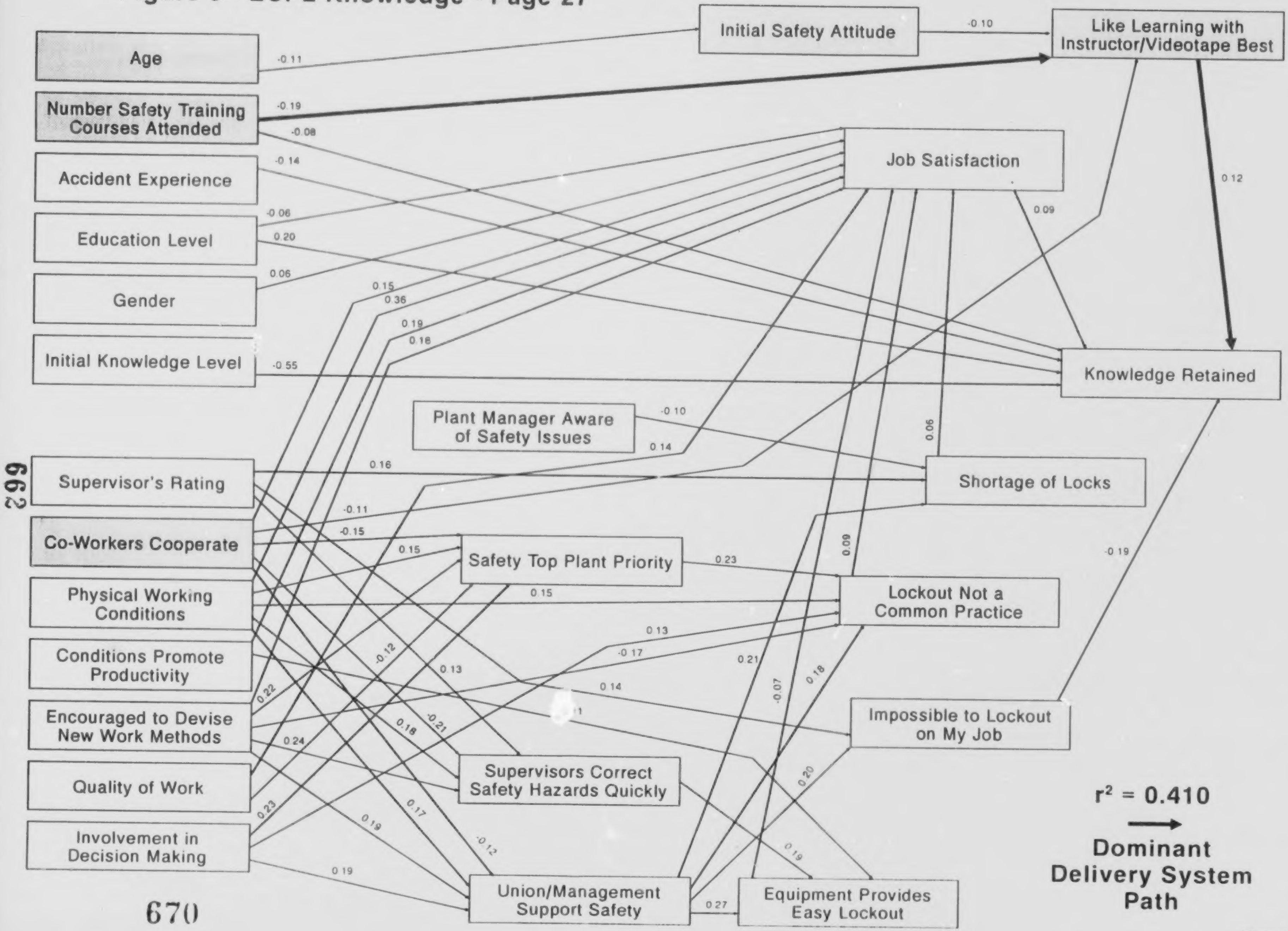
2. Gains in Attitude Towards Safety. The pre-test/post-test differences between the measure "The safety risks of my job concern me quite a bit."

3. Gains in General On-the-Job Safety Behavior. The pre-test/post-test differences between the measure "Before starting a job, how often do you consciously evaluate the consequences of not doing the job safely?"# and "How do you rate your safety performance?"##

APPENDIX B

PATH DIAGRAMS SUPPORTING GENERAL MODEL OF DELIVERY SYSTEM PREFERENCE EFFECTS

Figure 9 - ECPL Knowledge - Page 27



Age
Number Safety Training Courses Attended
Accident Experience
Education Level
Gender
Initial Knowledge Level
Initial Safety Attitude
Like Learning with Instructor/Videotape Best
Job Satisfaction
Knowledge Retained
Plant Manager Aware of Safety Issues
Shortage of Locks
Safety Top Plant Priority
Lockout Not a Common Practice
Impossible to Lockout on My Job
Supervisor's Rating
Co-Workers Cooperate
Physical Working Conditions
Conditions Promote Productivity
Encouraged to Devise New Work Methods
Quality of Work
Involvement in Decision Making
Supervisors Correct Safety Hazards Quickly
Union/Management Support Safety
Equipment Provides Easy Lockout
-0.11
-0.10
-0.19
-0.08
-0.14
-0.06
0.20
0.06
0.15
0.36
0.19
0.18
-0.55
0.12
0.09
-0.10
0.14
0.06
0.16
-0.11
-0.15
0.15
0.23
0.15
0.09
-0.19
0.13
-0.17
-0.12
0.13
0.14
0.21
-0.07
0.18
0.22
0.11
0.18
-0.21
0.24
0.20
0.17
0.19
0.23
-0.12
0.19
0.19
0.27
r² = 0.410
Dominant Delivery System Path

TABLE 5
Hypothesized Effects of Delivery System Preferences on Training Knowledge Retention
(ECPL Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Liked Learning with Instructor/ Videotape Best	Constant	0.768	0.067		
	Number of Programs	-0.025	0.007	-3.830	➡0.000+
	Rating of Self	-0.036	0.017	-2.189	➡0.029
	Mutual Safety Attitude	-0.054	0.028	-1.927	0.054
	$r^2 = 0.068$		N = 388		
Job Satisfaction Mean = 2.15 S.D. = 1.27	Constant	0.052	0.247		
	Physical Conditions	0.310	0.036	8.707	➡0.000+
	Conditions Promote	0.160	0.040	4.017	➡0.000+
	Co-workers Cooperate	0.147	0.040	3.683	➡0.000+
	Encouraged to Devise	0.187	0.040	4.690	➡0.000+
	Quality of Work	0.104	0.033	3.162	➡0.002
	Lockout Not Common	0.068	0.029	2.327	➡0.021
	Easy to Lockout	-0.091	0.038	-2.392	➡0.017
	Educational Level	-0.056	0.025	-2.231	➡0.026
	Gender	0.212	0.113	1.876	0.062
	Lockout Shortage	0.036	0.021	1.706	0.086
	Age	-0.064	0.039	-1.644	0.101
	$r^2 = 0.657$		N = 381		
Initial Safety Attitude Mean = 1.51 S.D. = 0.88	Constant	1.850	0.163		
	Age	-0.110	0.047	-2.149	➡0.032
	$r^2 = 0.012$		N = 381		
Impossible to Lockout on My Job Mean = 2.42 S.D. = 1.76	Constant	1.379	0.212		
	Union/Mgmt Support	0.322	0.081	3.956	➡0.000+
	Supervisor's Rating	0.177	0.065	2.719	➡0.007
	$r^2 = 0.070$		N = 388		
Lockout Not Common Practice Mean = 2.52 S.D. = 1.31	Constant	1.296	0.159		
	Safety Top Priority	0.238	0.059	4.074	➡0.000+
	Union/Mgmt Support	0.212	0.070	3.052	➡0.002
	Physical Conditions	0.149	0.053	2.642	➡0.009
	Encouraged to Devise	-0.162	0.058	-2.787	➡0.006
	Decision Involvement	0.112	0.050	2.270	➡0.024
	$r^2 = 0.190$		N = 388		

➡Significant at 0.05 level or less	➡Significant at 0.01 level or less

663 672

TABLE 5 (Continued)
Hypothesized Effects of Delivery System Preferences on Training Knowledge Retention
(ECPL Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Lock Shortage	Constant	2.017	0.226		
	Union/Mgmt Support	0.335	0.082	4.071	➨0.000+
	Supervisor's Rating	0.203	0.065	3.103	➨0.002
Mean = 2.94	Plant Manger Aware	-0.120	0.061	-1.970	➡0.050
S.D. = 1.77					
	r^2 = 0.082		N= 388		
Equipment Provides Easy Lockout	Constant	1.261	0.124		
	Union/Mgmt Support	0.252	0.049	5.113	➨0.000+
	Supervisors Correct	0.172	0.046	3.722	➨0.000+
	Conditions Promote	0.077	0.035	2.228	➡0.026
Mean = 2.21					
S.D. = 1.03	r^2 = 0.171		N = 388		
Union/ Management Support Safety	Constant	1.145	0.122		
	Decision Involvement	0.133	0.041	3.231	➨0.001
	Encouraged to Devise	0.146	0.051	2.867	➨0.004
	Physical Conditions	0.139	0.048	2.812	➨0.005
	Co-workers Cooperate	-0.099	0.050	-1.985	➡0.048
Mean = 1.96					
S.D. = 1.09					
	r^2 = 0.160		N = 388		
Supervisors Correct Hazards Quickly	Constant	1.893	0.133		
	Encouraged to Devise	0.198	0.052	3.787	➨0.000+
	Physical Conditions	0.147	0.051	2.878	➨0.004
	Co-workers Cooperate	-0.194	0.056	-3.454	➨0.000+
	Supervisor's Rating	0.107	0.049	2.194	➡0.029
Mean = 2.56					
S.D. = 1.16					
	r^2 = 0.125		N = 388		
	Constant	1.500	0.144		
Safety Top Plant Priority	Decision Involvement	0.187	0.049	3.842	➨0.000+
	Encouraged to Devise	0.199	0.060	3.332	➨0.000+
	Quality of Work	-0.110	0.053	-2.059	➡0.040
	Physical Conditions	0.135	0.056	2.397	➡0.017
Mean = 2.26	Co-workers Cooperate	-.0146	0.062	-2.333	➡0.020
S.D. = 1.27					
	r^2 = 0.159		N = 388		

➡Significant at 0.05 level or less | ➨Significant at 0.01 level or less

TABLE 5 (Continued)
Hypothesized Effects of Delivery System Preferences on Training Knowledge Retention
(ECPL Training)

Dependent Variable	Causal Variable	B	SE B	T	p
Gain in ECPL Knowledge Retained	Constant	15.199	1.638		
	Initial Knowledge Level	-1.038	0.088	-11.787	➡0.000+
	Educational Level	0.489	0.117	4.195	➡0.000+
	Lockout Impossible	-0.392	0.098	-3.977	➡0.000+
	Accident Experience	-0.528	0.180	-2.925	➡0.004
	Inst/Videotape Best	0.824	0.340	2.420	➨0.016
Mean = 2.48	Job Satisfaction	0.261	0.144	1.806	0.072
S.D. = 3.69	Number of Programs	-0.075	0.045	-1.680	0.094

$r^2 = 0.410$ N = 283

➨Significant at 0.05 level or less	➡Significant at 0.01 level or less

FIGURE 10 - Pedestrian Knowledge - Page 31

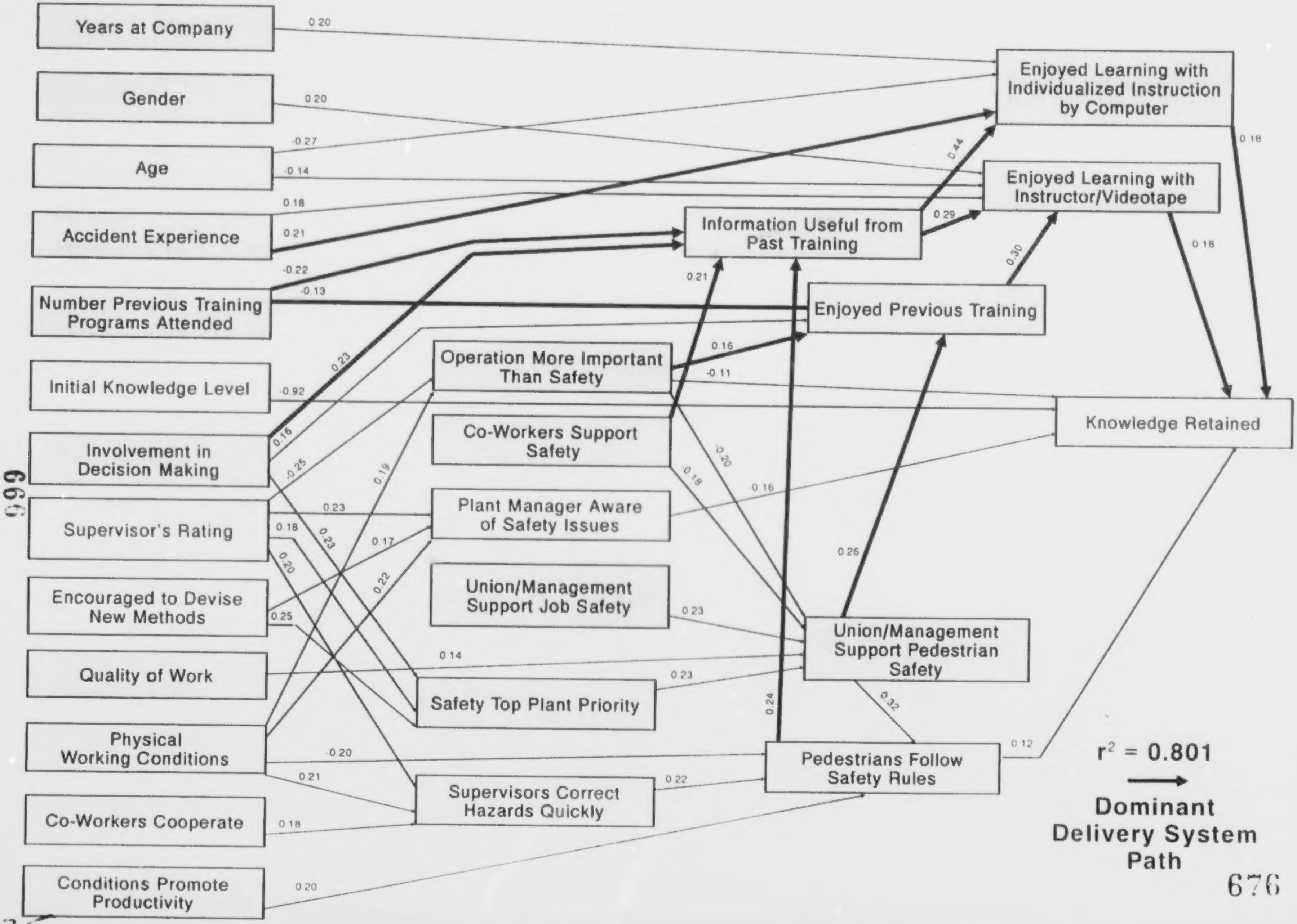

Years at Company
Gender
Age
Accident Experience
Number Previous Training Programs Attended
Initial Knowledge Level
Involvement in Decision Making
Supervisor's Rating
Encouraged to Devise New Methods
Quality of Work
Physical Working Conditions
Co-Workers Cooperate
Conditions Promote Productivity
Operation More Important Than Safety
Co-Workers Support Safety
Plant Manager Aware of Safety Issues
Union/Management Support Job Safety
Safety Top Plant Priority
Supervisors Correct Hazards Quickly
Information Useful from Past Training
Enjoyed Previous Training
Union/Management Support Pedestrian Safety
Pedestrians Follow Safety Rules
Enjoyed Learning with Individualized Instruction by Computer
Enjoyed Learning with Instructor/Videotape
Knowledge Retained
0.20
0.20
-0.27
-0.14
0.18
0.21
-0.22
-0.13
0.23
-0.92
0.16
-0.25
0.19
0.23
0.18
0.23
0.17
0.20
0.22
0.25
0.14
-0.20
0.21
0.18
0.20
0.44
0.29
0.30
0.21
0.16
-0.11
-0.20
-0.18
-0.16
0.26
0.23
0.23
0.24
0.32
0.22
0.12
0.18
0.18
$r^2 = 0.801$
Dominant Delivery System Path

TABLE 6
Hypothesized Effects of Delivery System Preferences on Training Knowledge Retention
(PMHV Pedestrian Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Enjoyed Learning with Instructor/ Videotape	Constant	0.269	0.354		
	Past Tng Info Useful	0.352	0.107	3.297	➠0.001
	Enjoyed Past Training	0.366	0.108	3.401	➠0.001
	Accident Experience	0.150	0.060	2.509	➡0.013
	Gender	-0.130	0.065	-2.011	➡0.046
Mean = 2.06	r^2 = 0.369		N = 138		
Enjoyed Learning with Individualized Instruction by Computer	Constant	1.304	0.440		
	Past Tng Info Useful	0.689	0.127	5.416	➠0.000+
	Accident Experience	0.229	0.090	2.538	➡0.013
	Age	-0.306	0.120	-2.547	➡0.012
	Years at Company	0.133	0.070	1.886	0.062
Mean = 2.65	r^2 = 0.295		N = 112		
Union/Mgmt Support Pedestrian Safety Rules	Constant	1.059	0.259		
	Safety Top Plant Priority	0.160	0.048	3.355	➠0.001
	Union/Mgmt Support	0.192	0.056	3.418	➠0.000+
	Operation Important	-0.153	0.050	-3.053	➠0.003
	Co-workers Support	0.165	0.055	2.974	➠0.003
Mean = 2.12 S.D. = 0.87	Quality of Work	0.126	0.058	2.178	➡0.030
	r^2 = 0.365		N = 187		
Safety Top Plant Priority	Constant	0.591	0.247		
	Decision Involvement	0.272	0.089	3.041	➠0.003
	Encouraged to Devise	0.298	0.084	3.543	➠0.001
Mean = 2.50 S.D. = 1.26	Supervisor's Rating	0.207	0.080	2.604	➠0.010
	r^2 = 0.265		N = 194		
Supervisors Correct Hazards Quickly	Constant	0.616	0.247		
	Supervisor's Rating	0.406	0.071	5.761	➠0.000+
	Physical Conditions	0.249	0.080	3.126	➠0.002
	Co-workers Cooperate	0.223	0.076	2.927	➠0.004
Mean = 2.82 S.D. = 1.21	r^2 = 0.324		N = 194		

➡Significant at 0.05 level or less | ➠Significant at 0.01 level or less

TABLE 6 (Continued)
Hypothesized Effects of Delivery System Preferences on Training Knowledge Retention
(PMHV Pedestrian Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Plant Mgr Aware of Safety Issues	Constant	0.785	0.256		
	Supervisor's Rating	0.233	0.075	3.106	➡0.002
	Physical Conditions	0.252	0.081	3.110	➡0.002
	Encouraged to Devise	0.191	0.075	2.527	→0.012
Mean = 2.46 S.D. = 1.15	r^2 = 0.205		N = 190		
Enjoyed Past Training Programs	Constant	1.626	0.230		
	Programs Attended	-0.130	0.030	-4.340	➡0.000+
	Union/Mgmt Support	0.199	0.060	3.290	➡0.001
	Decision Involvement	0.101	0.045	2.239	→0.026
	Operation Important	0.092	0.044	2.098	→0.037
Mean = 2.12 S.D. = 0.69	r^2 = 0.169		N = 179		
Past Training Information Useful	Constant	1.279	0.190		
	Decision Involvement	0.142	0.044	3.219	➡0.002
	Co-workers Support	0.153	0.051	3.015	➡0.003
	Pedestrians Follow	0.145	0.044	3.298	➡0.001
	Number of Programs	-0.093	0.030	-3.120	➡0.002
Mean = 2.03 S.D. = 0.70	r^2 = 0.247		N = 164		
Pedestrians Follow Safety Rules	Constant	1.293	0.247		
	Union/Mgmt Support	0.388	0.082	4.715	➡0.000+
	Supervisors Correct	0.192	0.064	2.993	➡0.003
	Conditions Promote	0.200	0.071	2.792	➡0.006
	Physical Conditions	-0.212	0.077	-2.744	➡0.007
Mean = 2.58 S.D. = 1.06	r^2 = 0.240		N = 187		
Operation More Important	Constant	3.522	0.239		
	Supervisor's Rating	-0.258	0.076	-3.413	➡0.001
	Physical Conditions	-0.218	0.083	-2.635	➡0.009
Mean = 2.33 S.D. = 1.16	r^2 = 0.134		N = 192		

→Significant at 0.05 level or less

➡Significant at 0.01 level or less

FIGURE 11 - Pedestrian Attitude - Page 35

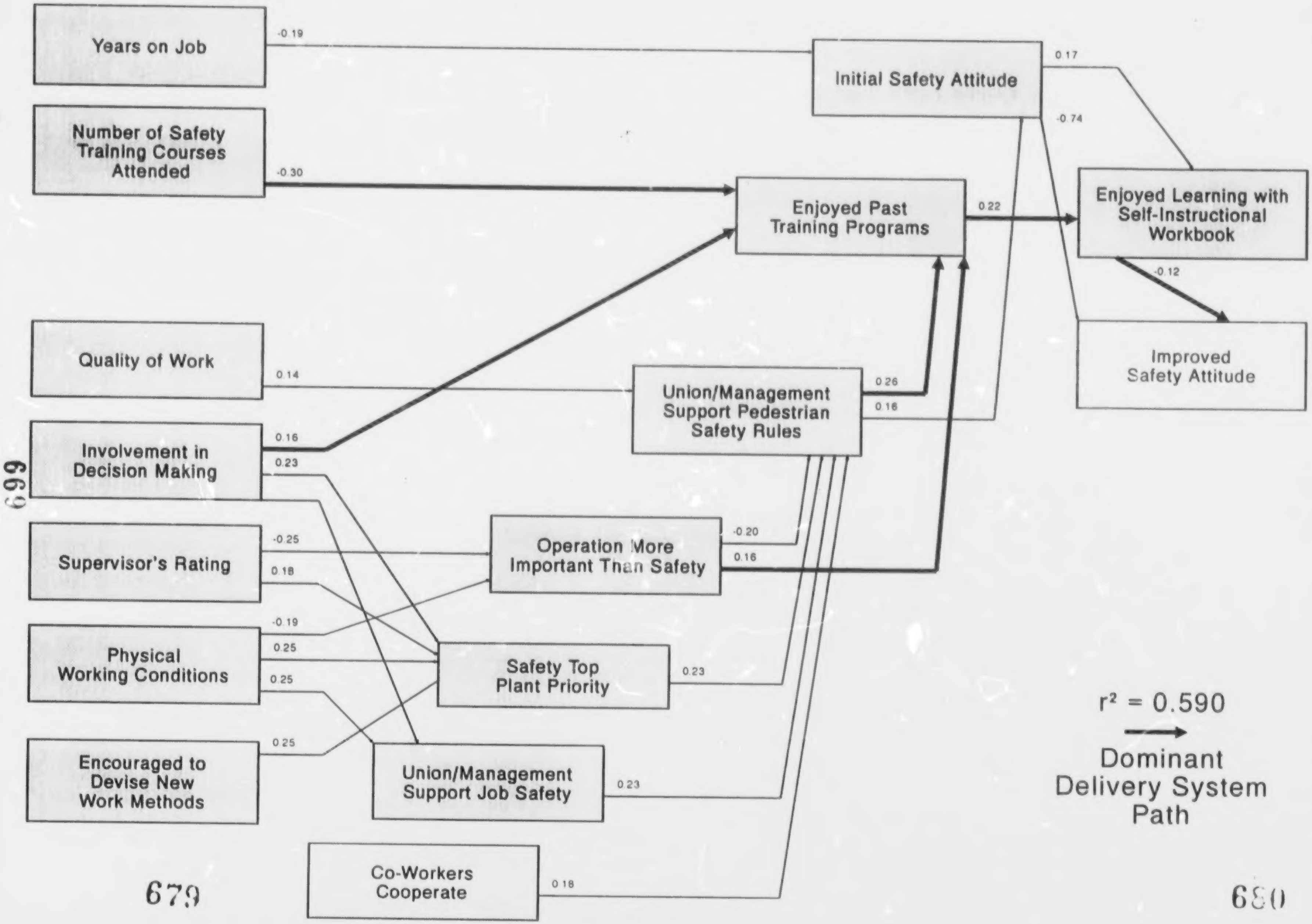

Years on Job
-0.19
Initial Safety Attitude
0.17
-0.74
Number of Safety Training Courses Attended
-0.30
Enjoyed Past Training Programs
0.22
Enjoyed Learning with Self-Instructional Workbook
-0.12
Improved Safety Attitude
Quality of Work
0.14
Union/Management Support Pedestrian Safety Rules
0.26
0.16
Involvement in Decision Making
0.16
0.23
Supervisor's Rating
-0.25
0.18
Operation More Important Than Safety
-0.20
0.16
Physical Working Conditions
-0.19
0.25
0.25
Safety Top Plant Priority
0.23
Encouraged to Devise New Work Methods
0.25
Union/Management Support Job Safety
0.23
Co-Workers Cooperate
0.18
r² = 0.590
Dominant Delivery System Path

TABLE 7
Hypothesized Effects of Delivery System Preferences on Attitude Changes After Training
(PMHV Pedestrian Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Enjoyed Learning with Self-Inst Workbook	Constant	1.738	0.259		
	Enjoyed Past Training	0.296	0.110	2.701	➠0.008
	Initial Safety Attitude	0.146	0.070	2.080	➡0.039
Mean = 2.57	$r^2 = 0.078$		N = 145		
Enjoyed Past Training Programs	Constant	1.626	0.230		
	Number of Programs	-0.130	0.030	-4.340	➠0.000+
	Union/Mgmt Support	0.199	0.060	3.290	➠0.001
	Decision Involvement	0.101	0.045	2.239	➡0.026
	Operation Important	0.092	0.044	2.098	➡0.037
Mean = 2.12 *S.D. = 0.693*	$r^2 = 0.169$		N = 179		
Initial Safety Attitude	Constant	1.545	0.207		
	Years on Job	-0.124	0.045	-2.731	➠0.007
	Union/Mgmt Support	0.177	0.079	2.248	➡0.026
Mean = 1.67 *S.D. = 0.98*	$r^2 = 0.066$		N = 191		
Union/ Management Support Pedestrian Safety Rules	Constant	1.059	0.259		
	Safety Top Priority	0.160	0.048	3.355	➠0.001
	Union/Mgmt Support	0.192	0.056	3.418	➠0.000+
	Operation Important	-.0153	0.050	-3.053	➠0.003
	Co-workers Support	0.165	0.055	2.974	➠0.003
	Quality of Work	0.126	0.058	2.178	➡0.031
Mean = 2.12 *S.D. = 0.87*	$r^2 = 0.365$		N = 187		
Union/Mgmt Support Safety	Constant	0.867	0.210		
	Physical Conditions	0.261	0.074	3.514	➠0.000+
	Decision Involvement	0.239	0.070	3.433	➠0.000+
Mean = 2.14 *S.D. = 1.03*	$r^2 = 0.177$		N = 194		
Safety Top Plant Priority	Constant	0.591	0.247		
	Decision Involvement	0.272	0.089	3.041	➠0.003
	Encouraged to Devise	0.298	0.084	3.543	➠0.000+
	Supervisor's Rating	0.207	0.080	2.604	➠0.010
Mean = 2.50 *S.D. = 1.16*	$r^2 = 0.265$		N = 194		

➡Significant at 0.05 level or less

➠Significant at 0.01 level or less

TABLE 7 (Continued)
Hypothesized Effects of Delivery System Preferences on Attitude Changes After Training
(PMHV Pedestrian Safety Training)

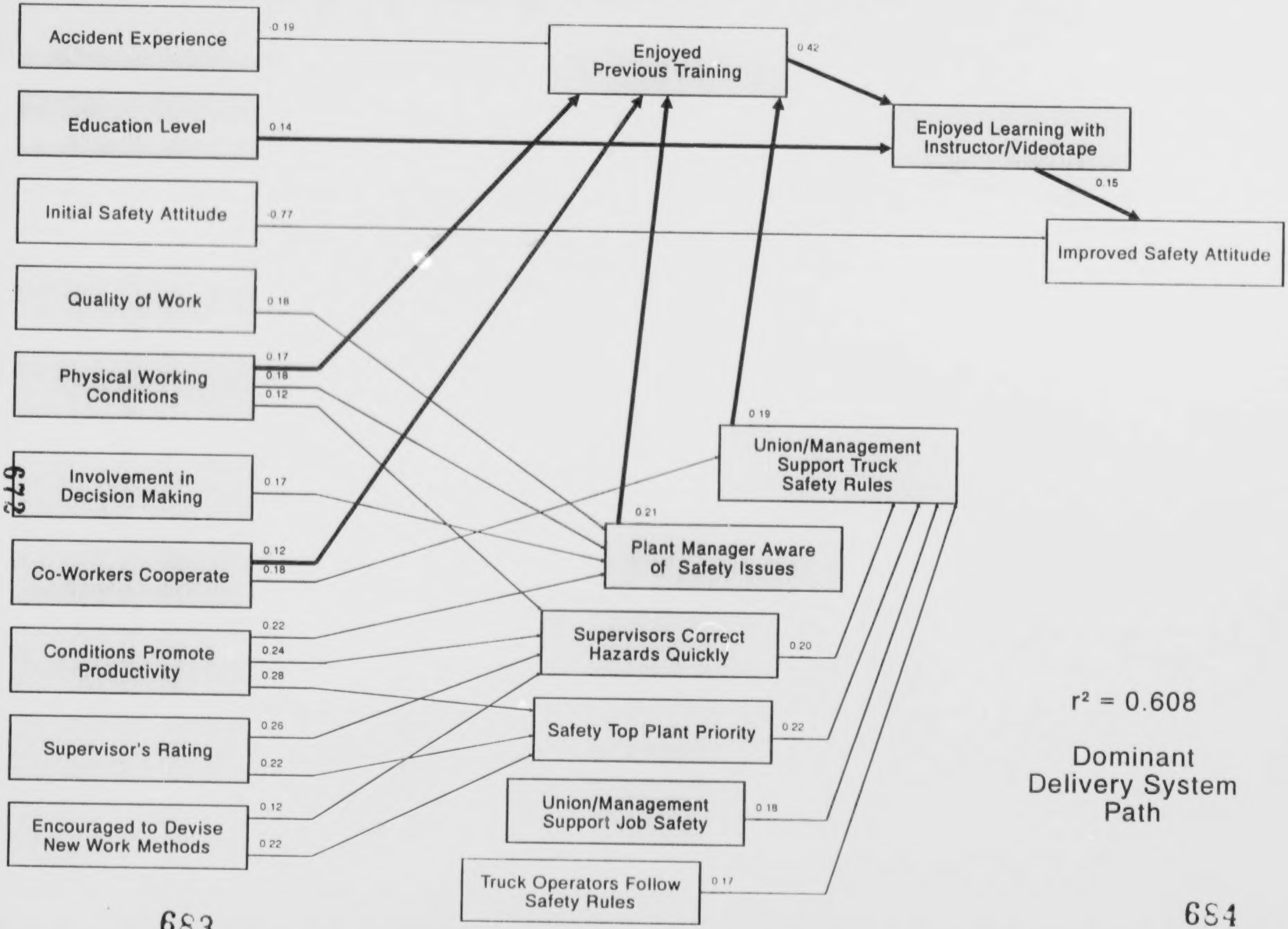
Endogenous Variable	Causal Variable	B	SE B	T	p
Operation More Important Than Safety	Constant	3.522	0.239		
	Supervisor's Rating	-0.258	0.076	-3.413	➡0.000+
	Physical Conditions	-0.218	0.083	-2.635	➡0.009
Mean = 2.33 *S.D. = 1.16*	$r^2 = 0.134$		N = 192		

Dependent Variable	Causal Variable	B	SE B	T	p
Gain in Safety Attitude	Constant	2.152	0.238		
	Initial Safety Attitude	-1.024	0.077	-13.260	
	Enjoy Self-Inst Wkbook	-0.165	0.080	-2.068	➥0.041
	$r^2 = 0.590$		N = 135		

➥Significant at 0.05 level or less

➡Significant at 0.01 level or less

FIGURE 12 - Operator Attitude - Page 38


Accident Experience
-0.19
Enjoyed Previous Training
0.42
Education Level
0.14
Enjoyed Learning with Instructor/Videotape
0.15
Initial Safety Attitude
-0.77
Improved Safety Attitude
Quality of Work
0.18
Physical Working Conditions
0.17
0.18
0.12
0.19
Union/Management Support Truck Safety Rules
Involvement in Decision Making
0.17
0.21
Plant Manager Aware of Safety Issues
Co-Workers Cooperate
0.12
0.18
Conditions Promote Productivity
0.22
0.24
0.28
Supervisors Correct Hazards Quickly
0.20
Supervisor's Rating
0.26
0.22
Safety Top Plant Priority
0.22
Encouraged to Devise New Work Methods
0.12
0.22
Union/Management Support Job Safety
0.18
Truck Operators Follow Safety Rules
0.17

$r^2 = 0.608$

Dominant
Delivery System
Path

TABLE 8
Hypothesized Effects of Delivery System Preferences on Attitude Changes After Training
(PMHV Operator Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Enjoyed Learning with Inst/Videotape	Constant	0.667	0.201		
	Enjoyed Past Training	0.471	0.064	7.329	➡0.000+
	Educational Level	0.148	0.061	2.416	0.016
Mean = 1.76	r^2 = 0.194		N = 252		
Enjoyed Past Training Programs	Constant	1.250	0.141		
	Plant Manager Aware	0.132	0.040	3.296	➡0.001
	Union/Mgmt Support	0.161	0.049	3.266	➡0.001
	Accident Experience	-0.148	0.042	-3.529	➡0.000+
	Physical Conditions	0.120	0.043	2.761	➡0.006
	Co-workers Cooperate	0.093	0.045	2.052	➥0.041
Mean = 2.07 S.D. = 0.72	r^2 = 0.230		N = 281		
Union/ Management Support Truck Safety Rules	Constant	0.362	0.130		
	Supervisor Corrects	0.147	0.046	3.219	➡0.001
	Safety Top Priority	0.170	0.046	3.721	➡0.000+
	Union/Mgmt Support	0.159	0.051	3.112	➡0.002
	Operators Follow	0.131	0.041	3.164	➡0.002
	Co-workers Cooperate	0.096	0.043	2.204	➥0.028
Mean = 1.93 S.D. = 0.89	r^2 = 0.401		N = 296		
Safety Top Plant Priority	Constant	0.489	0.160		
	Conditions Promote	0.275	0.057	4.846	➡0.000+
	Supervisor's Rating	0.272	0.064	4.266	➡0.000+
	Encouraged to Devise	0.218	0.059	3.705	➡0.000+
Mean = 2.21 S.D. = 1.11	r^2 = 0.302		N = 307		
Supervisorss Correct Hazards Quickly	Constant	0.606	0.175		
	Conditions Promote	0.239	0.062	3.844	➡0.000+
	Supervisor's Rating	0.337	0.068	4.940	➡0.000+
	Physical Conditions	0.132	0.064	2.081	➥0.038
	Encouraged to Devise	0.124	0.061	2.030	➥0.043
Mean = 2.46 S.D. = 1.13	r^2 = 0.297		N = 306		

➥Significant at 0.05 level or less ➡Significant at 0.01 level or less

TABLE 8 (Continued)
Hypothesized Effects of Delivery System Preferences on Attitude Changes After Training
(PMHV Operator Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Plant Manager Aware of Safety Issues	Constant	0.488	0.178		
	Physical Conditions	0.201	0.067	3.015	➡0.003
	Conditions Promote	0.220	0.057	3.884	➡0.000+
	Quality of Work	0.236	0.072	3.280	➡0.001
	Decision Involvement	0.172	0.058	2.966	➡0.003
Mean = 2.38 *S.D. = 1.12*		$r^2 = 0.300$		N = 301	

Dependent Variable	Causal Variable	B	SE B	T	p
Gains in Safety Attitude	Constant	1.188	0.145		
	Initial Safety Attitude	-1.059	0.062	-17.216	➡0.000+
	Enjoy Inst/Videotape	0.180	0.054	3.353	➡0.001
		$r^2 = 0.608$		N = 199	

➡Significant at 0.05 level or less	➡Significant at 0.01 level or less

FIGURE 13 - Pedestrian Behavior - Page 41

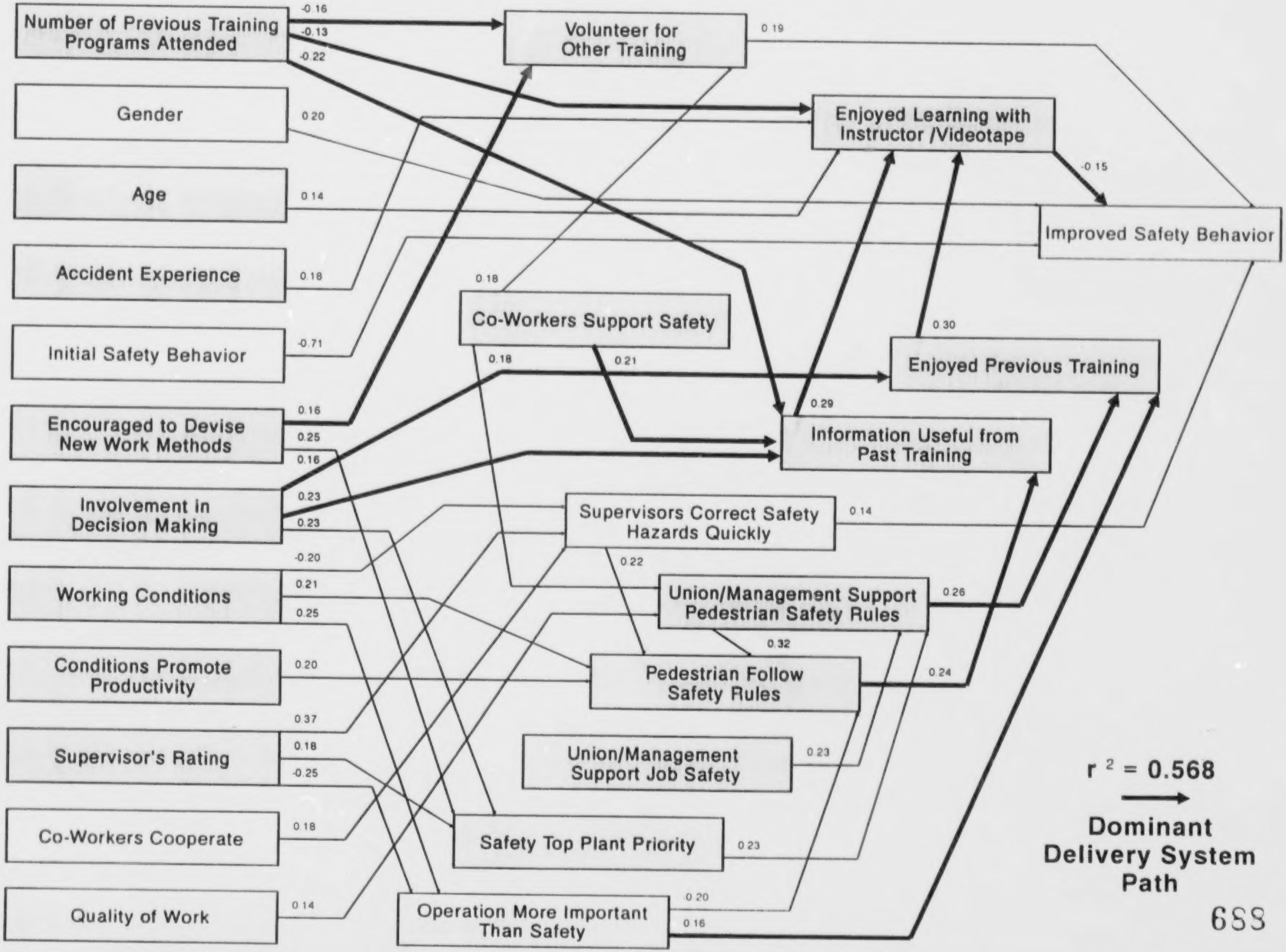

Number of Previous Training Programs Attended
-0.16
-0.13
-0.22
Volunteer for Other Training
0.19
Gender
0.20
Enjoyed Learning with Instructor /Videotape
-0.15
Age
0.14
Improved Safety Behavior
Accident Experience
0.18
0.18
Co-Workers Support Safety
0.30
Initial Safety Behavior
-0.71
0.18
0.21
Enjoyed Previous Training
Encouraged to Devise New Work Methods
0.16
0.25
0.16
0.29
Information Useful from Past Training
Involvement in Decision Making
0.23
0.23
Supervisors Correct Safety Hazards Quickly
0.14
-0.20
0.22
Working Conditions
0.21
0.25
Union/Management Support Pedestrian Safety Rules
0.26
0.32
Conditions Promote Productivity
0.20
Pedestrian Follow Safety Rules
0.24
0.37
Union/Management Support Job Safety
0.23
Supervisor's Rating
0.18
-0.25
Co-Workers Cooperate
0.18
Safety Top Plant Priority
0.23
Quality of Work
0.14
Operation More Important Than Safety
0.20
0.16
r 2 = 0.568
Dominant Delivery System Path

TABLE 9
Hypothesized Effects of Delivery System Preferences on General Behavior Changes After Training (PMHV Pedestrian Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Enjoyed Learning with Instructor/ Videotape	Constant	0.269	0.354		
	Past Tng Info Useful	0.352	0.107	3.297	➡0.001
	Enjoyed Past Training	0.366	0.108	3.401	➡0.001
	Accident Experience	0.150	0.060	2.509	➨0.013
	Gender	-0.130	0.065	-2.011	➨0.046
Mean = 2.06	$r^2 = 0.369$		N = 138		
Enjoyed Past Training Programs	Constant	1.626	0.230		
	Programs Attended	-0.130	0.030	-4.340	➡0.000+
	Union/Mgmt Support	0.199	0.060	3.290	➡0.001
	Decision Involvement	0.101	0.045	2.239	➨0.026
	Operation Important	0.092	0.044	2.098	➨0.037
Mean = 2.12 S.D. = 0.69	$r^2 = 0.169$		N = 179		
Past Training Information Useful	Constant	1.279	0.190		
	Decision Involvement	0.142	0.044	3.219	➡0.002
	Co-Workers Support	0.153	0.051	3.015	➡0.003
	Pedestrians Follow	0.145	0.044	3.298	➡0.001
	Number of Programs	-0.093	0.030	-3.120	➡0.002
Mean = 2.03 S.D. = 0.70	$r^2 = 0.247$		N = 164		
Volunteer for Other Training	Constant	1.882	0.238		
	Co-workers Support	0.166	0.067	2.475	➨0.014
	Number of Programs	-0.089	0.040	-2.207	➨0.029
	Encouraged to Devise	0.129	0.059	2.268	➨0.032
Mean = 2.22 S.D. = 0.89	$r^2 = 0.089$		N = 183		
Pedestrians Follow Safety Rules	Constant	1.293	0.247		
	Union/Mgmt Support	0.388	0.082	4.715	➡0.000+
	Supervisors Correct	0.192	0.064	2.993	➡0.003
	Conditions Promote	0.200	0.071	2.792	➡0.006
	Physical Conditions	-0.212	0.077	-2.744	➡0.007
Mean = 2.58 S.D. = 1.06	$r^2 = 0.240$		N = 187		

➨Significant at 0.05 level or less

➡Significant at 0.01 level or less

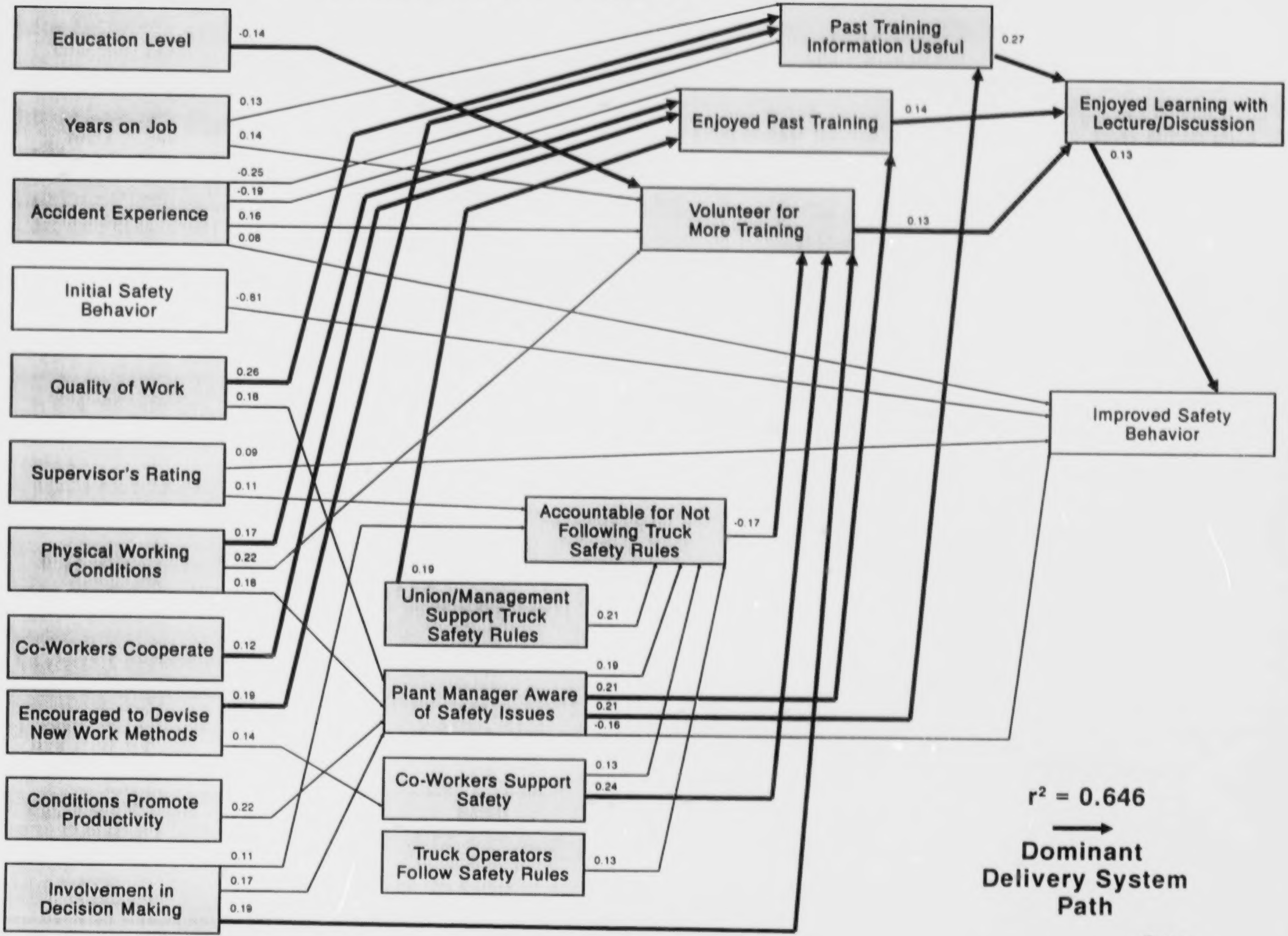
TABLE 9 (Continued)
Hypothesized Effects of Delivery System Preferences on General Behavior Changes After Training
(PMHV Pedestrian Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Union/Mgmt Support Pedestrian Safety Rules Mean = 2.12 S.D. = 0.87	Constant	1.059	0.259		
	Safety Top Plant Priority	0.160	0.048	3.355	➡0.001
	Union/Mgmt Support	0.192	0.056	3.418	➡0.000+
	Operation Important	-0.153	0.050	-3.053	➡0.003
	Co-workers Support	0.165	0.055	2.974	➡0.003
	Quality of Work	0.126	0.058	2.178	➨0.030
	$r^2 = 0.365$		N = 187		
Operation More Important Mean = 2.33 S.D. = 1.16	Constant	3.522	0.239		
	Supervisor's Rating	-0.258	0.076	-3.413	➡0.001
	Physical Conditions	-0.218	0.083	-2.635	➡0.009
	$r^2 = 0.134$		N = 192		
Safety Top Plant Priority Mean = 2.50 S.D. = 1.26	Constant	0.591	0.247		
	Decision Involvement	0.272	0.089	3.041	➡0.003
	Encouraged to Devise	0.298	0.084	3.543	➡0.001
	Supervisor's Rating	0.207	0.080	2.604	➡0.010
	$r^2 = 0.265$		N = 194		
Supervisors Correct Hazards Quickly Mean = 2.82 S.D. = 1.21	Constant	0.616	0.247		
	Supervisor's Rating	0.406	0.071	5.761	➡0.000+
	Physical Conditions	0.249	0.080	3.126	➡0.002
	Co-workers Cooperate	0.223	0.076	2.927	➡0.004
	$r^2 = 0.324$		N = 194		

Dependent Variable	Causal Variable	B	SE B	T	p
Gain in General Safety Behavior	Constant	1.430	0.362		
	Initial Safety Behavior	-0.963	0.079	-12.237	➡0.000+
	Volunteer for Programs	0.309	0.097	3.185	➡0.002
	Enjoy Inst/Videotape	-0.265	0.105	-2.515	➨0.013
	Supervisors Correct	0.164	0.068	2.405	➨0.018
	$r^2 = 0.568$		N = 134		

➨Significant at 0.05 level or less	➡Significant at 0.01 level or less

FIGURE 14 - Operator Behavior - Page 44



Education Level

Years on Job

Accident Experience

Initial Safety Behavior

Quality of Work

Supervisor's Rating

Physical Working Conditions

Co-Workers Cooperate

Encouraged to Devise New Work Methods

Conditions Promote Productivity

Involvement in Decision Making

Past Training Information Useful

Enjoyed Past Training

Volunteer for More Training

Enjoyed Learning with Lecture/Discussion

Improved Safety Behavior

Accountable for Not Following Truck Safety Rules

Union/Management Support Truck Safety Rules

Plant Manager Aware of Safety Issues

Co-Workers Support Safety

Truck Operators Follow Safety Rules

-0.14
0.13
0.14
-0.25
-0.19
0.16
0.08
-0.61
0.26
0.18
0.09
0.11
0.17
0.22
0.18
0.12
0.19
0.14
0.22
0.11
0.17
0.19
0.27
0.14
0.13
0.13
-0.17
0.19
0.21
0.19
0.21
0.21
-0.16
0.13
0.24
0.13

$r^2 = 0.646$

→ Dominant Delivery System Path

TABLE 10
Hypothesized Effects of Delivery System Preferences on General Behavior Changes After Training (PMHV Operator Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Enjoyed Learning with Lecture/ Discussion	Constant	0.845	0.141		
	Past Tng Info Useful	0.290	0.080	3.596	➡0.000+
	Volunteer for Training	0.102	0.052	1.960	0.051
	Enjoyed Past Training	0.140	0.076	1.846	0.066
Mean - 1.91					
	r^2 - 0.205		N - 255		
	Constant	1.582	0.270		
	Physical Conditions	0.203	0.064	3.189	➡0.002
Volunteer for Other Training	Co-workers Support	0.228	0.062	3.700	➡0.000+
	Educational Level	-0.173	0.078	-2.218	→0.028
	Decision Involvement	0.164	0.058	2.802	➡0.006
	Accountable for Rules	-0.159	0.065	-2.461	→0.015
Mean - 2.09	Accident Experience	-0.159	0.066	-2.388	→0.018
S.D. - 0.93	Years on Job	0.082	0.037	2.166	→0.031
	r^2 - 0.199		N - 214		
	Constant	1.217	0.162		
Past Training Information Useful	Quality of Work	0.217	0.054	4.028	➡0.000+
	Plant Manager Aware	0.132	0.042	3.123	➡0.002
	Accident Experience	-0.202	0.049	-4.089	➡0.000+
	Encouraged to Devise	0.124	0.042	2.946	➡0.004
Mean - 2.04	Years on Job	0.059	0.028	2.121	→0.035
S.D. - 0.71					
	r^2 - 0.260		N - 212		
	Constant	1.250	0.141		
Enjoyed Past Training Programs	Plant Manager Aware	0.132	0.040	3.296	➡0.001
	Union/Mgmt Support	0.161	0.049	3.266	➡0.001
	Accident Experience	-0.148	0.042	-3.529	➡0.000+
	Physical Conditions	0.120	0.043	2.761	➡0.006
Mean - 2.07	Co-workers Cooperate	0.093	0.045	2.052	→0.041
S.D. - 0.72					
	r^2 - 0.230		N - 281		
Co-workers Support Safety	Constant	0.544	0.129		
	Encouraged to Devise	0.132	0.053	2.507	→0.013
Mean - 1.83					
S.D. - 1.00	r^2 - 0.020		N - 305		

→Significant at 0.05 level or less

➡Significant at 0.01 level or less

TABLE 10 (Continued)
Hypothesized Effects of Delivery System Preferences on General Behavior Changes After Training
(PMHV Operator Safety Training)

Endogenous Variable	Causal Variable	B	SE B	T	p
Accountable For Not Following Truck Safety Rules	Constant	0.148	0.171		
	Union/Mgmt Support	0.239	0.065	3.658	➡0.000+
	Plant Manager Aware	0.164	0.052	3.168	➡0.002
	Supervisor's Rating	0.115	0.059	1.953	0.052
	Co-workers Support	0.121	0.048	2.496	➡0.013
	Operators Follow Rules	0.110	0.051	2.159	➡0.031
	Decision Involvement	0.010	0.048	2.082	➡0.038
Mean = 1.95 S.D. = 0.954	r^2 = 0.321		N = 291		
Union/ Management Support Truck Safety Rules	Constant	0.362	0.130		
	Supervisor Corrects	0.147	0.046	3.219	➡0.001
	Safety Top Priority	0.170	0.046	3.721	➡0.000+
	Union/Mgmt Support	0.159	0.051	3.112	➡0.002
	Operators Follow	0.131	0.041	3.164	➡0.002
	Co-workers Cooperate	0.096	0.043	2.204	➡0.028
Mean = 1.93 S.D. = 0.89	r^2 = 0.401		N = 296		
Plant Manager Aware of Safety Issues	Constant	0.488	0.178		
	Physical Conditions	0.201	0.067	3.015	➡0.003
	Conditions Promote	0.220	0.057	3.884	➡0.000+
	Quality of Work	0.236	0.072	3.280	➡0.001
	Decision Involvement	0.172	0.058	2.966	➡0.003
Mean = 2.38 S.D. = 1.12	r^2 = 0.300		N = 301		
Dependent Variable	**Causal Variable**	**B**	**SE B**	**T**	**p**
Gain in General Safety Behavior	Constant	1.324	0.175		
	Initial Safety Behavior	-1.017	0.064	-15.955	➡0.000+
	Enjoy Lecture/Discussion	0.166	0.058	2.841	➡0.005
	Plant Manager Aware	-0.132	0.041	-3.250	➡0.001
	Accident Experience	0.083	0.047	1.759	0.080
	Supervisor's Rating	0.097	0.055	1.776	0.078
	r^2 = 0.646		N = 184		

Significant at 0.05 level or less	Significant at 0.01 level or less

Title:

The Effects of Visual Grouping on Learning from Computer Animated Presentations

Author:

Lloyd P. Rieber

The Effects of Visual Grouping on Learning from Computer Animated Presentations

Lloyd P. Rieber
Educational Technology Program
Texas A&M University

Abstract

The effects of visual grouping strategies involving animated and static graphic presentations on learning were studied. Also studied was the ability of students to learn a scientific rule presented incidentally in an animated sequence in the hope of replicating results from previous research. A total of 39 fourth graders participated in an introductory lesson on Newton's laws of motion. Two levels of Visual Presentation (Static Graphic, Animated Graphic) were crossed with two levels of Visual Grouping (Grouped, Ungrouped). A within-subjects factor consisted of two levels of Learning Intent (Intentional, Incidental). Results showed that students given animated presentations of lesson content outperformed students receiving static presentations, but only when the animated lesson frames were presented in groups, or "chunks," of textual and visual sequences. These results demonstrate how students' selective attention toward information contained in animated visual can be undermined when the presentation insufficiently cues their focus of attention to the animated sequence. These results also illustrate at least one possible strategy which can be used to improve the ability of students to attend to and learn from animated visuals. Results also showed that students were successfully able to extract information pertaining to an application of Newton's second law incidentally presented in animated sequences. These results successfully replicate earlier findings indicating that elementary school students can learn and apply incidental information from an animated display without necessarily sacrificing intentional learning.

Dramatic visual effects can be produced with computer animation. Intuition would suggest that learning from apparently well-designed animated displays would be equally dramatic. Unfortunately, research has not shown this to be the case. It appears that animated instruction, even in its best form, exerts a relatively subtle influence on learning and that many factors can serve to undermine its effectiveness (see review by Rieber, 1990-a). For example, students may have difficulty attending to relevant details contained in animated visual displays which may be obvious to the designer or an expert (Reed, 1985; White, 1984). However, several very recent reports have begun to clarify some of the conditions under which animation may facilitate learning (Rieber, 1990-b; Rieber, in press).

An example of the evasive nature of animation's effect on learning is provided by Rieber (1989) who reported no significant differences on the learning of science concepts between animated and static visual presentations. However, the elementary school subjects took significantly less *time* to process lesson frames containing the animated sequences than similar frames containing static graphics or all text. It was speculated that students did not effectively attend to the animation or may have been distracted by the combination of visual and verbal information presented to them. An even more likely hypothesis is that students used the time needed to display the animation for other mental tasks, such as reading screen text, thereby obviating any benefit to learning which the animation might have provided. From a theoretical perspective, this is the problem of selective attention (Woolfolk, 1987). Animation is probably an effective attention-gaining device, especially for brief periods, but when animation is designed to be informational or instructive, learning demands sustained meaningful effort on the part of students. Students must be able to perceive important information and also actively select and process this information in the proper context. This is especially important given the temporal nature of animation — students may not have a second chance to view the sequence once it has ended.

The Rieber (1989) study was replicated with a variety of changes to the lesson's instructional design. One such change involved modifying the way in which the lesson frames were presented to students. Instead of presenting frames one at a time, as conventional design wisdom suggests, each frame was broken down and presented in a series of three or four textual and visual portions, or "chunks." It was believed that this change would help to direct students' attention to the information contained in the animated sequences by reducing the threat of distraction from other screen information, such as novel text. Results of the study showed differentiated learning effects for lessons contained animation as compared to those containing static visuals or no visuals (Rieber, 1990-b). There is a need to validate whether this modification actually accounted for differentiated learning from animated displays by helping to increase students' selective attention to lesson information contained in the display. The first objective of this study was to investigate this issue.

In addition to evidence that animation can aid intentional learning, further research suggests that animation may also promote incidental learning. In a recent study, students effectively extracted incidental information about a specific application of Newton's second law contained in animated display and were subsequently able to use this information to solve problems (Rieber, in press). The study of incidental learning, though well-established in the general literature (see Klauer, 1984, for a review), is usually overlooked in the design of instruction based on systems approaches (Branson & Grow, 1987). Proponents of incidental learning assert the ability of the learner to elaborate upon the given instruction and to induce concepts and principles from the context. Past studies of incidental learning have typically indicated tradeoffs between intentional and incidental learning. Faced with allocating limited cognitive processing abilities, students typically make gains on one kind of learning (e.g. intentional) at the expense of other kinds of learning (e.g. incidental). However, in the study by Rieber (in press), students were able to learn incidental information without any apparent sacrifices to intentional learning. The economical nature of spatial information in working memory may provide access to greater amounts of information when lesson information is presented in both verbal and visual forms (Gagné, 1985). One application of this finding, if it proves to be robust, is using incidental information learned in one lesson as an organizer for subsequent lessons where this information would be intentionally taught. The issue of whether this incidental learning result is replicable is questioned and is the basis for the second objective of this study.

In summary, there were two purposes to this study. First, the effects of visual grouping of lesson presentations containing animated and static visuals on learning were investigated. It was hypothesized that animated graphics presentations would facilitate learning more than static graphics, but only when presented in a grouped or "chunked" manner as this should significantly influence students to selectively attend to the relevant details contained in the animation. A secondary purpose was to replicate earlier findings that animated graphics promote incidental learning. It was hypothesized that students in lessons containing animated presentations would significantly outperform other students on measures of incidental learning as has been previously documented (Rieber, in press).

Method

Subjects

The subjects consisted of 39 fourth grade students from an urban public elementary school. Participation was voluntary with selection based on parent consent.

CBI Lesson Content

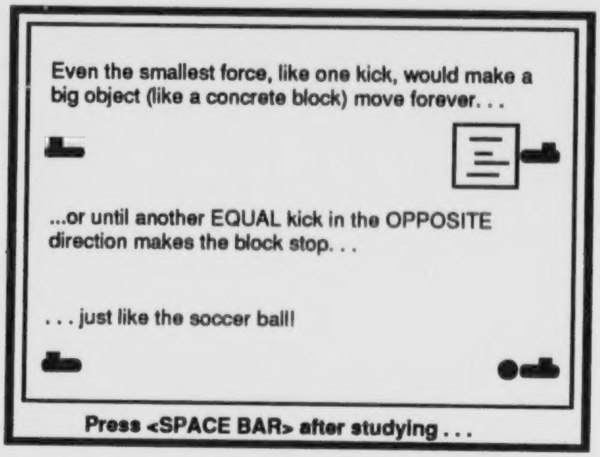
The CBI lesson, first developed for previous studies (Rieber, 1989; Rieber, 1990-b; Rieber, in press) but modified for the purposes of this study, described and explained Newton's laws of motion. The lesson was divided into two parts. Part one introduced students to Isaac Newton and his first law of motion (i.e. an object in motion tends to remain in motion and an object at rest tends to stay at rest). Part two provided applications of Newton's first law and introductions and applications of Newton's second law (i.e. force equals mass times acceleration). The lesson was designed and presented at an introductory level in the context of one-dimensional motion without friction and gravity.

Lesson Versions

Two levels of Visual Presentation (Static Graphic, Animated Graphic) were crossed with two levels of Visual Grouping (Grouped, Ungrouped).

Visual Presentation. In each Visual Presentation condition the lesson narrative was supplemented with either static or animated graphics. In the Static Graphics condition, forces were represented by a foot giving an object a kick, such as a ball or a concrete block. Motion and trajectory attributes were represented by arrows and path lines. The Animated Graphics condition also used the ball, block and foot symbols. However, the object was then animated to show the motion and trajectory resulting from the force applied to it. Both the static and animated visuals elaborated the textual information.

In addition to the information which was intentionally taught in both the static and animated graphic conditions, another application of Newton's second law — that dealing with the motion which results when the mass of an object varies — was *incidentally* demonstrated in the Animated Graphics condition. This law predicts that when two objects of different masses are both acted on by the same size force, the acceleration experienced by the object with larger mass is less. The computer animated a large "concrete block" moving very slowly across the computer screen after being kicked with the same size force which had just been



Even the smallest force, like one kick, would make a big object (like a concrete block) move forever. . .

...or until another EQUAL kick in the OPPOSITE direction makes the block stop. . .

. . . just like the soccer ball!

Press <SPACE BAR> after studying . . .

FIGURE 1. *Example of a computer screen which used animation to present intentional and incidental information. The "concrete block" and the "ball" were both kicked once. The animation depicted the ball's speed as dramatically greater than that of the block. Copyright 1991 by the American Psychological Association.*

applied to a small animated "soccer ball," similar to what is illustrated in Figure 1. The incidental information contained in the animation attributes was in the form of varying the *speed* of the two moving objects of different masses on the computer screen. No attempt was made to teach or explain why the block moved more slowly than the ball. By design, no incidental information was contained in the static visuals as no information related to "speed" was contained in the arrows or path lines.

***Visual Grouping*.** The verbal (textual) and visual information in each lesson frame was presented to subjects in one of two ways: grouped and ungrouped. Each frame of instruction in each grouping condition contained exactly the same information formatted in the same way. The difference between the two groups was the way in which the information was presented to the subjects. In the ungrouped condition, as in most CBI, lesson information was presented to subjects one frame at a time as one continuous unit. Students were prompted to press the space bar when ready to view the next frame. In the grouped condition, each frame was presented to subjects as a series of three of four information groups, or "chunks," in order to cue subjects' attention to the animated visuals and to minimize screen distractions. Each subject was then prompted to press the space bar when ready to view the next chunk, as illustrated in Figure 2.

Dependent Measures

***Posttest*.** The 24-item posttest measured two types of rule learning: intentional and incidental. Intentional learning is defined as learning those objectives which were directly taught in the lesson through the combination of verbal (text) and visual means. The intentional learning outcomes included applications of Newton's first law and a specific application of Newton's second law in which the mass of an object is held constant and the forces which act upon it are allowed to vary. Incidental learning is defined as learning those objectives which were not directly taught, but only implied through contextual cues provided in several of the animated visual displays. The incidental learning outcome was another application of Newton's second law in which the forces acting on objects are held constant and the mass of the objects is allowed to vary. Half of the intentional and incidental questions were presented in an all-verbal form and half contained accompanying visuals.

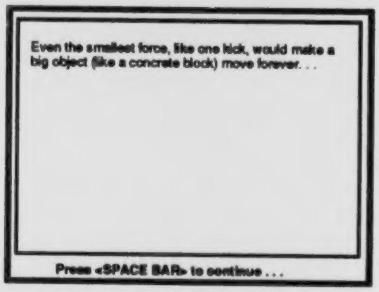
KR-20 reliability of all 24 questions was .86. Respective reliability coefficients and number of questions for each of the subscales were as follows: .78, 14 (Intentional); .90, 10 (Incidental).

Procedure

All instruction and testing was administered by computer. As students arrived at the computer lab they were randomly assigned to one of the treatment groups. After a group orientation to the lesson, students completed the lesson individually. Approximately 75 minutes was needed by students to complete the lesson and posttest.

Design

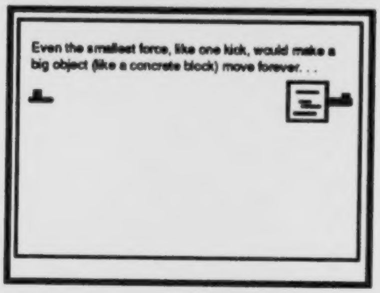
In this study a 2 X 2 X (2) factorial design was used. Two between-subjects factors (Visual Presentation,



Even the smallest force, like one kick, would make a big object (like a concrete block) move forever. . .

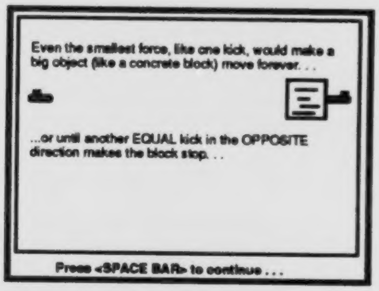
Press <SPACE BAR> to continue . . .

Frame chunk1: Text



Even the smallest force, like one kick, would make a big object (like a concrete block) move forever. . .

Frame chunk 2A: Animation sequence immediately followed by ...



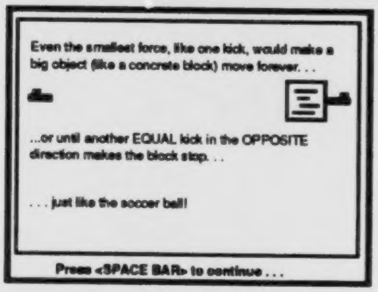
Even the smallest force, like one kick, would make a big object (like a concrete block) move forever. . .

...or until another EQUAL kick in the OPPOSITE direction makes the block stop. . .

Press <SPACE BAR> to continue . . .

... frame chunk 2B: Text

FIGURE 2. *An example of how a single lesson frame was grouped into verbal and visual portions, or "chunks." Each student pressed the space bar when ready to view the next chunk.*



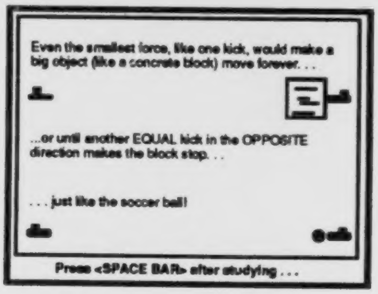
Even the smallest force, like one kick, would make a big object (like a concrete block) move forever. . .

...or until another EQUAL kick in the OPPOSITE direction makes the block stop. . .

. . . just like the soccer ball!

Press <SPACE BAR> to continue . . .

Frame chunk 3: Text



Even the smallest force, like one kick, would make a big object (like a concrete block) move forever. . .

...or until another EQUAL kick in the OPPOSITE direction makes the block stop. . .

. . . just like the soccer ball!

Press <SPACE BAR> after studying . . .

Frame chunk 4: Animation

TABLE 1. *Mean Percent Scores and Standard Deviations for the Posttest*

Visual Presentation	Visual Grouping		Learning Outcome: Intentional	Learning Outcome: Incidental
Animated Graphic	Grouped	Mean	79.3	69.0
	(n=10)	SD	17.0	34.1
	Ungrouped	Mean	59.3	54.0
	(n=10)	SD	29.7	31.3
Static Graphic	Grouped	Mean	66.7	17.8
	(n=9)	SD	22.6	22.8
	Ungrouped	Mean	71.4	32.0
	(n=10)	SD	16.8	30.1

Notes: SD-standard deviation; n-number of subjects

Visual Grouping) were crossed with one within-subjects factors (Learning Intent). The between-subjects factors consisted of two levels of Visual Presentation (Static Graphic, Animated Graphic) and two levels of Visual Grouping (Grouped, Ungrouped). The within-subjects factor consisted of two levels of Learning Intent (Intentional, Incidental). The within-subjects factor also served as the dependent measure. A mixed factor analysis of variance (ANOVA) was used to analyze the quantitative data.

Results

Percent means and standard deviations of student posttest scores are contained in Table 1. A significant interaction was found between Visual Presentation and Visual Grouping, $F(1,70)=5.08$, $p<.05$, $MS_e=697.71$. Follow-up analyses indicated that students in the Animated Grouped condition (mean=74.1%) performed significantly better on the posttest than students in either the Static Grouped (mean=42.2%) or Static Ungrouped (mean=51.7%) conditions, as shown in Figure 3. The Animated Ungrouped condition (mean=56.6%) was not significantly different from any of the other three conditions. These results support the first hypothesis of this study that students in the Animated Graphics condition would outperform those in the Static Graphics condition, but only when properly cued and guided to attend to the information contained in the animated display. Animated graphics presented continuously with other frame information were no more effective than either of the static graphics conditions. These results reflect the differences reported between the Rieber (1989) and Rieber (1990-b) studies.

A significant interaction was also found between Visual Presentation and Learning Intent, $F(1,70)=9.22$, $p<.01$, $MS_e=697.71$. Students in the Animated and Static Graphics conditions performed similarly on intentional learning (means=69.3% and 69.2%, respectively), however, students in the Animated Graphics condition (mean=61.5%) dramatically outperformed those in the static graphics condition (mean=25.3%) on incidental learning, as shown in Figure 4. As expected, these latter students performed at about the rate expected for guessing on incidental questions. This result supports the second hypothesis of this study which stated that students receiving animated demonstrations of incidental information would attend, understand, and retrieve this information when tested although no formal attempt was made to teach or explain the principle. This finding successfully replicates the incidental learning results reported by Rieber (in press).

FIGURE 3. *Interaction between Visual Presentation and Visual Grouping.*

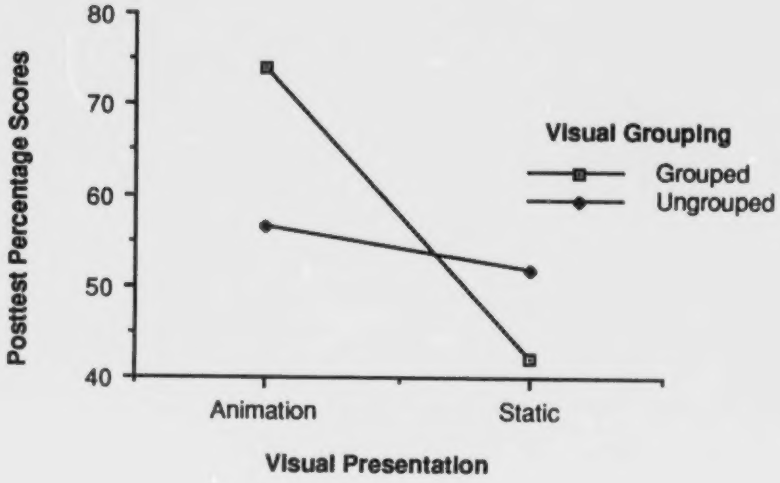
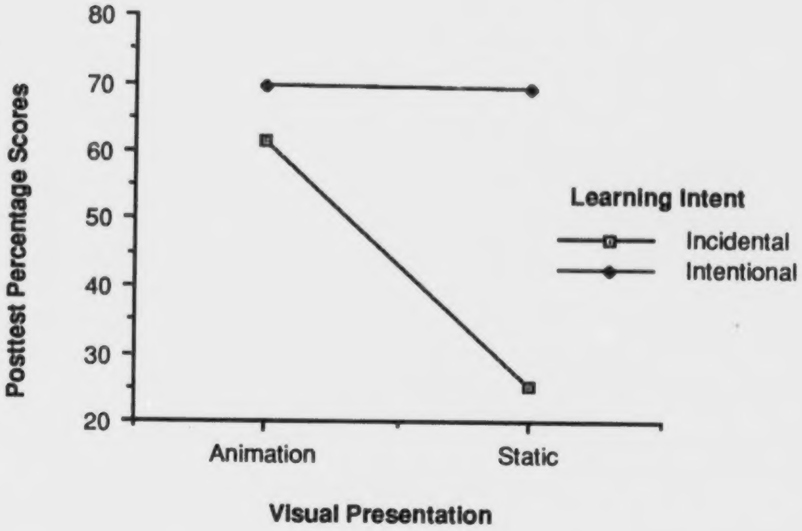


FIGURE 4. *Interaction between Visual Presentation and Learning Intent.*



Discussion

There were two purposes to this study — one primary and one secondary. The primary purpose was to investigate whether students' selective attention to animated information is influenced by the method in which it is presented. In this study, students in the Animated Grouped condition performed significantly better on the posttest than students in either of the Static Graphics conditions. In other words, animation was a significant aid to learning as compared to static graphics, but only when presented as part of a series of verbal and visual chunks. Students who were presented with lesson frames which contained static visuals were not influenced by the method of grouping.

These results were anticipated. Past research which has compared the use of animated and static visuals as adjunct presentation aids has indicated that animation's influence on learning, if any, are relatively subtle (e.g. Caraballo, 1985; King, 1975; Reed, 1985; Rigney & Lutz [Alesandrini], 1975). Perceptually, animated visuals offer two attributes beyond that of static visuals: motion and trajectory (i.e. direction of travel) (see Rieber, in press; and Rieber & Kini, 1991, for a more thorough description of the theoretical foundations of instructional applications of animated visuals). Therefore, learning from instructional tasks which require students to visualize the motion of objects along certain paths should be improved with appropriate applications of computer animation. It was hypothesized that animation would be effective when its attributes were congruent to the learning task *and* when presented to students in a manner that effectively cued their attention to the information contained in the animated visual. Past research has suggested that children do not necessarily know how to "read" an animated visual designed to illustrate an intellectual skill, perhaps because much of the animation they experience is meant to appeal to their affective domain, such as cartoons. In this study, a simple grouping strategy was used in the presentation of the computer lesson frames. Each frame was broken down into a series of three of four verbal and visual portions, or "chunks," which were presented individually to students. Students in these Grouped conditions had control over the pacing of these lesson frame chunks. The results of this study support the hypothesis that students attention to animation must be sufficiently cued and directed in order for potential learning effects derived from animation to be realized. It is suggested that the grouping strategy used here is but one example of a possible attention focusing device that might be used.

Adjunct static graphics would not be expected to be influenced by the grouping strategy used here given the fact that students had as much time as they needed to view each lesson frame. Since a static graphic, by definition, does not change over time, students had sufficient opportunity to study these visuals and their relationship to the lesson text. Animated visuals, on the other hand, demand temporal considerations. If students do not educe the information they contain, either because they did not understand the message or did not attend to the sequence, they may not have another opportunity to view the sequence and correct their learning deficiency, as was the case in this study. Many alternative strategies which help students attend to the information undoubtedly exist beyond the simple grouping strategy used here. For example, simply providing students with the opportunity to repeat an animated sequence may be effective, assuming, of course, that students are monitoring their comprehension sufficiently to know that they did not comprehend the animated sequence and that the option to repeat exists.

Much of what is known about instructional visuals comes from the pool of static visual research. Recent reviews suggest a variety of guidelines and conditions under which static visuals contribute to learning (see Alesandrini, 1984; Dwyer, 1978; Kobayashi, 1986; Levie & Lentz, 1982; Levin, Anglin, & Carney, 1987). Among the most important groundrules guiding the use of visuals is congruency between presentation and task variables (Levin & Lesgold, 1978). It is suggested that animated visuals represent a subset of instructional visuals and therefore represent extensions of the empirical and theoretical base of static visuals (such as that provided in the dual-coding hypothesis offered by Paivio, 1986). The results of this study suggest an additional important groundrule for the successful application of animated visuals in instruction: students' attention must be adequately cued to the information contained in an animated visuals and special consideration needs to be taken to avoid having their attention distracted away from the animation, even if the animation is well-designed in all other ways, otherwise no learning may occur.

The second goal of this study was to replicate earlier findings that elementary school students are able to extract incidental information contained in an animated visual. Besides providing a visual elaboration of the intentional learning outcomes, the animated sequences also demonstrated incidentally an application of Newton's second law where the masses of two objects varied, but the forces acting on them remained constant. The incidental information was represented by the relative motion of screen objects: an animated "concrete block" moved obviously slower than a "soccer ball" after equal size forces were applied to each. The students in the Animated Graphics conditions were apparently able to attend to, learn, and apply a sophisticated science principle without any apparent sacrifice to intentional learning. These results successfully replicate similar and earlier findings (Rieber, in press).

In conclusion, this study has demonstrated the fragile influence that animated visuals exert on learning. Learning from apparently well-designed animated visuals can be subverted when students' attention is not sharply focused on the relevant information contained in an animated sequence. Designers and developers must be careful not to confuse the dramatic visual effects of animation with its related effects on learning. Young children especially appear to need external prompting or cueing in order to effectively learn from animated visuals. However, learning can occur when students are focused on an animated visual. The replicated findings that students can learn and apply a scientific rule presented incidentally by an animated visuals offer a glimpse of the potential of animation in learning environments. The incidental information contained in the animation was simple and easily discernible by students, such that they did not even need explicit guidance to appropriately learn and apply the information. Several directions for future research and developmental work are suggested. More research is needed to clarify and validate the grouping strategy used in this study as well as other potentially effective cueing strategies. Much development work is needed to appropriately exploit the incidental learning gains suggested by the use of animated visuals, such as using this incidental learning as a bridge to intentional learning in subsequent lessons.

References

Alesandrini, K. (1984). Pictures and adult learning. *Instructional Science, 13*, 63-77.

Branson, R., & Grow, G. (1987). Instructional systems development. In R. Gagné (Ed.), *Instructional Technology: Foundations* (pp. 397-428). Hillsdale, NJ: Lawrence Erlbaum Associates.

Caraballo, A. (1985). *An experimental study to investigate the effects of computer animation on the understanding and retention of selected levels of learning outcomes.* Unpublished doctoral dissertation, The Pennsylvania State University.

Dwyer, F. (1978). *Strategies for improving visual learning.* State College, PA: Learning Services.

Gagné, E. (1985). *The cognitive psychology of school learning.* Boston: Little, Brown, and Company.

King, W. (1975). *A comparison of three combinations of text and graphics for concept learning.* (Report No. NPRDC-TR-76-16). San Diego, CA: Navy Personnel Research and Development Center. (ERIC Document Reproduction Service No. ED 112 936)

Klauer, K. (1984). Intentional and incidental learning with instructional texts: A meta-analysis for 1970-1980. *American Educational Research Journal, 21*, 323-339.

Kobayashi, S. (1986). Theoretical issues concerning superiority of pictures over words and sentences in memory. *Perceptual and Motor Skills, 63*, 783-792.

Levie, W., & Lentz, R. (1982). Effects of text illustrations: A review of research. *Educational Communications and Technology Journal, 30*(4), 195-232.

Levin, J., Anglin, G., & Carney, R. (1987). On empirically validating functions of pictures in prose. In D. Willows & H. Houghton (Eds.), *The psychology of illustration, volume 2: Instructional issues* (pp. 51-85). New York: Springer-Verlag.

Levin, J., & Lesgold, A. (1978). On pictures in prose. *Educational Communications and Technology Journal, 26*(3), 233-243.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Reed, S. (1985). Effect of computer graphics on improving estimates to algebra word problems. *Journal of Educational Psychology, 77*(3), 285-298.

Rieber, L. (1989). The effects of computer animated elaboration strategies and practice on factual and application learning in an elementary science lesson. *Journal of Educational Computing Research, 5*(4), 431-444.

Rieber, L. (1990-a). Animation in computer-based instruction. *Educational Technology Research & Development, 38*(1), 77-86.

Rieber, L. (1990-b). Using computer animated graphics in science instruction with children. *Journal of Educational Psychology, 82*, 135-140.

Rieber, L. (in press). Animation, incidental learning, and continuing motivation. *Journal of Educational Psychology.*

Rieber, L., & Kini, A. (1991). *Theoretical foundations of instructional applications of computer-generated animated visuals.* Submitted for publication.

Rigney, J. & Lutz [Alesandrini], K. (1975). *The effects of interactive graphic analogies on recall of concepts in chemistry* (Technical Report No. 75). Washington, DC: Office of Naval Research. (ERIC Document Service No. ED 109 639)

White, B. (1984). Designing computer games to help physics students understand Newton's laws of motion. *Cognition and Instruction, 1*, 69-108.

Woolfolk, A. (1987). *Educational psychology* (3rd edition). Englewood Cliffs, NJ: Prentice-Hall.

Title:

Computer-based Microworlds: A Bridge between Constructivism and Direct Instruction

Author:

Lloyd P. Rieber

Computer-Based Microworlds: A Bridge Between Constructivism and Direct Instruction

Lloyd P. Rieber
Educational Technology Program
Texas A&M University

Abstract

The field of instructional technology is characterized by its products, such as instructional media, and its processes, such as instructional design. The process of instructional technology has been shaped by advances in learning and instructional theory over the past 50 years. Much of the development work of the field to date has been direct instruction, or instruction based largely on the application of behavioral principles. Most computer applications are based on principles of programmed instruction — the computer becoming the ultimate teaching machine. In contrast to behavioral learning, constructivism, a faction within cognitive psychology associated with the application of Piagetian learning theory, is characterized by discovery and experiential learning. Constructivists have sought to tap the computational power of modern microcomputers to create computer "microworlds" where learners can experience and appropriate sophisticated ideas from (but not limited to) the domains of science and mathematics. Probably the most well-known computer application of constructivism is LOGO. Proponents of constructivism and direct instruction are usually viewed in opposition to one another. It is suggested that computer microworlds offer an immediate compromise between these two camps. In this paper, constructivism and microworlds, as well as related research on mental models, misconceptions in science, and intrinsic motivation are reviewed. Finally, some considerations are offered in the design of microworlds which can be used an an aid to direct instruction. These design considerations are presented in the context of Space Shuttle Commander, *a prototype project developed to teach motion concepts based on compromises between principles of constructivism and direct instruction.*

Instructional technology can be defined as the application of what is known about learning and teaching on current instructional practices (Knirk & Gustafson, 1986). Instructional technology is by its very nature a practical concern and instructional technologists are essentially educational pragmatists. All share the mission of applying whatever knowledge bases are available to achieve "the goals" of the instructional system. Interestingly, instructional technology is often equated with one of many instructional media or models of instructional design such as those based on the instructional systems development (ISD) approach, but these are mere instances of instructional technology which do not cover its breadth of scope and purpose any more than addition or subtraction defines mathematics or engineering. Instructional technology is an evolving and dynamic field which is naturally susceptible to the tension between the many internal forces and perspectives, both current and historical, which seek to define it.

Learning theory is arguably the most significant foundation of instructional technology (Gagné & Glaser, 1987). Advances in learning theory have had considerable influence on instructional technology. Instructional technology, in its modern form, began as an attempt to apply behavioral learning principles to instruction and quickly merged with the audio-visual movement of the middle 1900's (see Reiser, 1987, for a historical review). Instructional technology has since proceeded to a neo-behavioral philosophy (Case & Bereiter, 1984) which considers the learner as a more active agent in the learning process. Most applications of computer-based learning have been extensions of the programmed instruction paradigm as set forth by Skinner (1968). The pioneering work of Patrick Suppes in the development of computer assisted instruction (CAI) in arithmetic is a prime example (Suppes & Morningstar, 1972). Cognitive psychology has become a major influence on instructional technology over the last 15 years, though probably more conceptually in the writings of the field than in actual practice (Gagné & Dick, 1983; Hannafin & Rieber, 1989; Shuell, 1986).

The ISD approach to instructional design has characterized much of instructional technology since World War II (Knirk & Gustafson, 1986). While many people still equate ISD with behavioral learning, it is argued that this presents too narrow a definition of current applications of ISD. Many ISD approaches and methods have tried to blend behaviorism and cognitivism in order to achieve learner-centered instruction which is still goal-oriented. Many are based on accretion or reception learning models where meaningful learning is seen as a progression through a series of stages along a novice to expert continuum (Ausubel, 1968; Norman, 1982; Mayer, 1984) or on a hierarchy from lower-level learning to higher-level learning (eg. Gagné, 1985). Expert status is achieved through a process where knowledge is successively acquired and organized, and then integrated or "fine tuned" into existing knowledge structures. One might label this approach to instructional technology as "instructivism" because the instructional goals are still the dominant influence on instructional design, despite the concern and care given to learner abilities and needs.

Among the achievements of cognitive psychology, the writings of Jean Piaget remain an important watershed. Though his work has been extended and more fully elaborated, Piaget's basic model of human learning and development remains intact. Piaget's work, though popular and influential in Europe for decades, was not embraced in this country until about 1970, having been overshadowed by the behaviorist movement until that time. Recent applications of Piaget's work over the last ten years are usually grouped under the heading of "constructivism." Interestingly, the influence of constructivism has only recently begun to spill over to the field of instructional technology. Probably the most well known computer-based application of constructivism is LOGO, a computer language designed to reflect and promote Piagetian learning. LOGO was the result of a collaborative effort between the Massachusetts Institute of Technology (MIT) and Bolt, Beranek, and Newman (BBN), and was initially funded by the National Science Foundation. Though LOGO involved the contributions of many people, such as Wally Feurzeig, Daniel Bobrow, Hal Abelson, and Andy diSessa, Seymour Papert is usually credited as LOGO's chief developer. LOGO is a primary example of an application of constructivism based on "microworlds." A microworld, as the name suggests, is a small, but complete subset of reality in which one can go to learn about a specific domain through personal discovery and exploration (Dede, 1987; Papert, 1981).

The major premise of this paper is that instructional technology in general, and educational computing in particular, need to seriously consider the infusion of constructivism into instructional design and that microworlds represent an immediate application. Microworlds can offer goal-oriented learning environments in which learning is achieved through activities which promote discovery and exploration. The compromise is reached largely through a *guided-discovery* orientation to learning where the nature of the learning activity and experience is naturally constrained by the parameters imposed by a particular microworld. Microworlds offer a compromise between the strict deductive approach suggested by ISD models and pure inductive learning advocated by experiential learning theorists. However, because the microworld's boundaries are severely limited, the range of learning outcomes is also constrained making the achievement of predetermined learning objectives possible and probable, which is the aim of instructivism.

In this paper, the constructivist position on learning based on several milestones of Piaget's theories will be summarized and applied to computer microworlds. Also reviewed will be two important areas which have a direct bearing on the design of microworlds for specific goals: intrinsic motivation and mental models. Intrinsic motivation is one of the important assumptions of the constructivist approach, but yet is rarely described in detail by its advocates. Mental models comprise an important area of research in cognitive psychology. Mental model research is helpful in understanding human knowledge about the world and subsequent human interaction within it. Most of the microworlds created thus far, like most of the mental model research conducted to date, have been in technical areas, such as physics and mathematics. Insights gained by mental model research closely parallels the goals of microworld design. Finally, a series of design principles will be presented which attempt to synthesize the collaboration of instructivism and constructivism. An original computer-based development project will be used in the presentation of these principles.

Constructivism

Constructivism is the name of a school of educational philosophy closely aligned with the theories of Jean Piaget (see Forman & Pufall, 1988a, for a review). At the heart of constructivism is the idea that learning involves individual constructions of knowledge. Constructivist proponents range from those who stress a strong humanistic orientation where learning occurs through open discovery (eg. Rogers, 1969) to those who accept the imposition of varying degrees of structure depending on the setting (eg. Howe, O'Shea, & Plane, 1980). Constructivists describe learning as occurring through interactions with one's environment or "culture." Therefore, the potential for learning at different levels is thought to grow as the environment becomes richer and more engaging. Computer enthusiasts, such as Papert (1980), feel that the computer offers a powerful medium for the exploration and discovery for many ideas, particularly in science and mathematics, just as a young child might explore the concepts of volume given a sandbox and mass and momentum given a bag of marbles.

Piaget's theory of cognitive development essentially has two parts: stage dependent and stage independent (Mayer, 1983). Most of the attention to Piaget's theory is usually given to the stage dependent part, or that which refers to the four stages of development that people presumably progress through: sensorimotor, preoperational, concrete operations, and formal operations. However, Piaget's stage independent theory is of most interest in the design of computer microworlds. It concerns two processes: the *adaptation* of an individual to survive and flourish in an ever changing environment; and the need for *organization*, or stability and coherence of the world knowledge and experiences learned "to date." Piaget referred to this organized set of knowledge and understanding as internal structures or schemata. The processes of adaptation and organization create an intrinsic conflict which provides the platform for lifelong learning. The goals of one process are in contrast to the goals of the other. Life entails a continual balancing between the two which is accomplished through the process of *equilibration*. *Assimilation* and *accommodation* are the two well-known mechanisms which carry out the process of equilibration. Assimilation subsumes new information under an already understood structure or schemata. A baby who has learned to throw a tennis ball is also likely to throw a rubber ball as well as an orange. Both the rubber ball and the orange fit the structure that one "throws round objects." Accommodation builds new structures from old structures when incoming new information no longer fits an existing structure. The baby soon learns to accommodate the fact that some round objects are to be thrown, and others are to be eaten.

Assimilation and accommodation result from the natural conflict which occurs through everyday encounters with the environment. Learning can only occur when an individual is in a state of disequilibrium, or cognitive conflict. Given new information, an individually intrinsically seeks to assimilate, or incorporate, the information into preexisting structures. Accommodation is triggered when new information no longer fits established structures, necessitating the adaption of new structures.

Constructivism, the instructional application of Piaget's work, consists of three properties: epistemic conflict, self-reflection, and self-regulation (Forman & Pufall, 1988b). Epistemic conflict is the Piagetian process of equilibration. Learning occurs when the environment presents a problem which is outside of the individual's repertoire but which the individual still seeks to resolve. Even though the conflict may be induced by some external agent, such as a teacher, resolution is achieved only by the individual. Each resolution is an individual construction. Self-reflection involves an individual's attempt at objectively and explicitly representing reality and is a response to conflict. Self-regulation is the spontaneous restructuring of thought which is embodied in the Piagetian process of accommodation and assimilation. Old mental structures are refined to be more comprehensive as well as new structures formed. Conflict and reflection are necessarily prerequisite to self-regulation.

Proponents of constructivism emphasize the quality of knowledge structures over their quantity. Learning is viewed not so much as the acquisition of knowledge, but as the constant reconstruction of what is already known. Individuals do not simply add information to their knowledge "banks," but either revise existing mental structures to accept new information or formulate new structures based on old ones when an existing structure is no longer sufficient. As Forman and Pufall (1988b) note: "Central to constructivism is the assumption that to know is to continually reconstruct, to move from a more to a less intuitive state" (p. 240).

One of the most promising attempts at creating computer environments which foster the construction (assimilation and accommodation) of knowledge are microworlds. Microworlds offer instructional designers two key advantages. First, microworlds present learners with experiences *within* specific boundaries of a domain. Second, microworlds offer learners with "stepping stones" *between* interconnected ideas within the domain by allowing rudimentary ideas to first become established and then transformed into more sophisticated aspects of the domain.

Computer Microworlds

A microworld is small, but complete, subset of a specific environment or domain. Microworlds are often confused with simulations. Characteristics of the two can heavily overlap or even be synonymous as well as remain entirely distinct, depending on their design and most importantly on how they are used in a learning interaction. A microworld has two essential characteristics which might help make this point more clear. First, a microworld embodies the simplest model of a domain that is deemed accurate and appropriate by an expert. Second, it offers an initial point of entry which matches the user's cognitive state to allow fruitful interactions to take place. For example, Cuisenaire rods act as a microworld for many mathematical ideas. They embody the manipulation of a set of introductory mathematical ideas which are valid from any mathematician's point of view, but function at the level of the children who use them. Cuisenaire rods offer not a mathematical simulation, but permit *real* mathematics to take place. On the other hand, simulations can be designed which do not offer any significant difference from real-life experiences, such as the sophisticated flight simulators used for training by the military and most major airlines. These simulations would not be considered microworlds for most people because they are designed to represent as many of the variables and factors of the real experience as possible. It is important to remember that microworlds, like simulations, do not have to be computer generated. However, the computational and graphical power of computers provide a powerful arsenal to microworld designers.

Simulations can become microworlds with design changes. A computer flight simulation can easily be designed to permit manipulation and experience with only one part of the aircraft, such as the rudder. In this sense, the simulation becomes a "rudder microworld." Similarly, many microworlds can easily become simulations. Consider a mathematical microworld which involves the estimation of distances such as by using the LOGO command FORWARD to move the turtle from one point to another on the screen with as few commands as possible. This microworld becomes a "whale search" simulation simply by changing the turtle into an animated "boat" and the screen target into a "whale." The mathematical microworld has not changed, only the context.

Papert (1980) suggests that microworlds, like all powerful ideas, should fulfill four criteria. They should be simple, general, useful, and syntonic. Syntonic learning means "it goes together with." Syntonic learning suggests that learning is made up of connections, such as connecting new ideas to old. Syntonic learning means going from the "known to the unknown" which is central to every idea in cognitive psychology. These learning connections are enhanced in a microworld through learner control. This is in contrast to research on learner control of direct instruction which frequently suggests that learners are often poor judges of their own learning paths (Clark, 1980; Steinberg, 1977).

Microworlds rely on the natural tendency of a learner to seek equilibrium. Successful microworlds actually encourage learning conflicts to arise in order to activate the process of equilibration. It is only through the resolution of these conflicts that learning can take place. The trick is to structure the microworld so that the learner has an environment in which conflict resolution is within their grasp. Microworlds offer learners with the opportunity to exercise a cognitive or intellectual skill which they would be unable or unlikely to do so on their own either because there is no intrinsic reason to do so or because no sufficient tool is available with which to allow the learner to construct the experience.

Creating "Objects to Think With"

LOGO represents the most ambitious attempt to date at creating computer-based microworlds in which learners have ready access to mathematical ideas and heuristic problem-solving. Another is BOXER, a computer-based microworld which combines computational and text-processing tools in a programming environment based on interacting windows called "boxes" (diSessa & Abelson, 1986). Whereas BOXER is still under development and is not yet available for general application, LOGO has had widespread dissemination on most microcomputers over the last 12 years. Turtle geometry is one such LOGO microworld which gives learners access to geometric principles through interactive graphics. The LOGO turtle is a cybernetic animal that "lives" on the computer screen and follows a set of commands like forward, back, right, and left. As students "drive" the turtle around the screen it leaves a trail, thus allowing students to create graphics through the manipulation of mathematics (see Abelson & diSessa, 1981; Lockard, Abrams, & Many, 1990; and Lukas & Lukas, 1986, for further descriptions and discussions of turtle geometry). Research on LOGO is inconclusive (see review by Clements, 1985), perhaps because, as Papert (1987) contends, LOGO is meant to be part of a cultural influence on learning and considering its effects on learning in isolation is not defensible. The issue here is not the success of LOGO but the way in which LOGO was developed to achieve constructivist goals.

Successful LOGO learning experiences are founded on many key ideas, many associated with programming. For example, LOGO is procedural, which encourages "top-down" problem-solving in which a large problem can be broken down into smaller, more manageable "chunks." However, the turtle embodies the essence of constructivism by being an "object to think with" (Papert, 1980). Papert (1988) goes so far as to suggest that the turtle may represent a "mother structure" for differential geometry. Mother structures are the most fundamental ideas or concepts upon which cognition is built.

At the heart of constructivism is a search for good objects to think with. Almost anything can become a good object to think with: pots, pans, mudpies, blocks, legos, etc. Some are more flexible and generalizable to a variety of domains than others. Meaningful interaction with objects in the environment liberates and encourages the equilibrium process. The turtle is but one example of an object to think with that is made possible by computer technology. Papert contends that the computational power of the turtle makes certain ideas from the world of mathematics and problem-solving accessible to children that previously were considered as too "formal" or abstract to be understood. Several characteristics of the turtle in particular help make this possible: the turtle as a transitional object; and the turtle as an aid to "debugging."

Young children often begin using LOGO for self-guided learning within minutes of encountering the turtle. This is largely achieved by the role of the turtle as a transitional object between themselves and the computer. The turtle is body syntonic with a child in that both share two important characteristics: a position and a heading. This simple fact has powerful consequences. From the start, even a young child has something in common with the turtle. This commonality immediately provides the bridge to new ideas. Young children quickly anthropomorphize the turtle creating an ego syntonic relationship with the turtle. This encourages the Piagetian concept of decentering in which young children begin to see the world from several perspectives. The anthropomorphization of the turtle also encourages the articulation and communication of mathematical ideas. Children begin to acquire the "vocabulary" of turtle geometry through their communications with the turtle. Since the language of the turtle is LOGO and the language of LOGO is mathematics, children have a way in which to verbalize mathematics.

The second important characteristic of the turtle is as an aid to "debugging" which is the identification and correction of errors within a computer program. In most forms of instruction errors are something to be avoided: errors imply failure. In LOGO and constructivism, errors are a natural consequence of interactions. Errors, instead of having negative connotations, are useful so long as they provide a rich source of information for subsequent interactions. Errors are a natural part of all human endeavors. The informational feedback that errors provide is very potent especially when a learner has a strong commitment to the action that triggered the error (eg. Kulhavy, 1977). Error detection is made intuitively obvious in LOGO with the turtle's role as a graphical tool. The turtle's animated graphics provide almost instantaneous graphic feedback to a learner. This rapid turn-over of learner action and feedback encourages risk taking and hypothesis testing. The forming and testing of hypotheses through graphical feedback becomes more important as a learner's computer program becomes more complex.

Mental Models

Daily activities require all of us to interact with a complex environment. Many researchers and theorists suggest that people form mental models of the physical world and the objects and activities within it (see Gentner & Stevens, 1983, for a review). A mental model is an individual's conceptualization of a specific domain or "system." Students develop and use mental models to help explain and solve general classes of problems. Far from formal and static entities, mental models are loosely organized and forever changing as new interactions with the environment suggest adaptations. Mental model research has concentrated on technical domains, like physics or electricity, because they are far more normative and are more easily made explicit than most other domains, such as "parenting." Although researchers may use highly specific domains, theorists suggest that people form mental models of a large number of systems ranging from the kitchen stove to Newtonian mechanics (see Norman, 1988, for a discussion of ways in which people form mental models of everyday things).

Mental models serve explanatory and predictive purposes. Survival demands that we be able to predict everyday events with a high degree of success. The routine need to cross a street is a prime example. Beyond all of the perceptual requirements (such as estimating the width of the street and the speed of oncoming cars) is a need to understand the many "street" systems which operate together. Just a few of these systems include the workings of an automobile, traffic lights, and physics. Our understanding of each of these systems is crucial as we decide when is an appropriate time to cross the street as well as if we can casually stroll across or should attempt an "Olympic sprint." Any misunderstanding of one of these systems could be as deadly as any misjudgment of distance or speed. For example, consider your mental model of an electric "Walk" sign at an intersection and what it means when it begins to flash. Can you still initiate the crossing? What should those already part way across do? Each interpretation can have dramatically different consequences. Mental models of everyday things usually form through interactions with the environment. However some systems, such as the physical sciences, are difficult to understand through a wide range of random interactions. Microworlds offer a platform for appropriate conceptualizations of such systems to form and develop.

The application of mental models to the instructional design of microworlds involves the consideration of three things: the target system, the conceptual model of that target system, and the user's mental model of the target system (Norman, 1983). The target system is the actual system that a learner is trying to understand. For example, Newtonian mechanics or thermodynamics are typical target systems. A user's mental model describes that individual's conception of the target system; a person describes how the target system works and predicts its behavior based on their mental model of the target system. A user's mental model may or may not be an accurate reflection of the target system. For example, research has shown that many people hold a mental model of motion problems based on Aristotelian rather than Newtonian principles (Champagne, Klopfer, & Anderson, 1980; diSessa, 1982). Hence, many students reckon that a ball comes to a stop soon after being thrown because it "runs out" of force. A conceptual model is presented to a learner as an aid to understanding the target system. Conceptual models are usually invented by teachers, designers, or engineers.

For all practical purposes, a microworld is largely synonymous with an *interactive* conceptual model. It embodies the simplest working model of a system in which an individual can begin to understand the target system. A conceptual model can often be metaphorical to the target system, such as suggesting that the flow of electrons in a wire is "like water running through a pipe." In such cases conceptual models, like microworlds, offer a temporary "doorway" to a set of larger ideas. For example, Papert (1980) has recounted the way in which his fascination with gears as a young boy offered him with a beginning conceptual model of mathematical ratios and proportions. For Papert, gears became a personal microworld which helped concretize many abstract mathematical ideas for him. This does not imply that the gears metaphor should be introduced into the mathematics curriculum, but instead suggests how a microworld can provide an introductory conceptual transition into a learning domain.

Even as adults, we all have and form mental models of everyday objects and events, such as kitchen stoves, bathroom shower controls, and light switch panels (Norman, 1988). A particularly good example is the thermostat of a home furnace. Many people believe that the higher you set the thermostat, the faster your home gets heated, when in reality this is not the case. A personal example concerns an old black and white television set I have in my bedroom. When I turn on the set, its power quickly goes on and then off. If I tap the set, it will quickly turn on and off again. In order to get the set to turn on and stay on, I have to carefully tap the set until the "connection catches." I have learned that there are better places to tap the set than others. I have become particularly adept at how to turn the set on, and I frequently coach my wife and daughter on the "technique." Now, I know nothing about television repair or about electrical devices, yet I have developed my own "loose connection" theory, or mental model, to define the set's electrical problem (assuming it is an electrical problem). I formed and tested many hypotheses until finally arriving at the one I feel is accurate. It is important to note that this was done informally and naturally. So far, this mental model allows me to predict future instances of the set's performance and thereby governs my actions.

Mental model research is closely aligned with research on misconceptions in science (Eylon & Linn, 1988). Misconceptions in science have been documented in many specific domains, from physics and optics to mathematics and biology. Misconceptions about the laws of motion have been a favorite domain of study. Research has shown that students form strong personal theories, though naive, about the physical world which are based on intuitive, and frequently inaccurate, models (Champagne, Klopfer, & Anderson, 1980). For example, it is commonly viewed that students tend to hold an incorrect Aristotelian model of the physical world. Some proponents of this view, most notably McCloskey (1983), suggest that the instructional course of action is to force a theory change in students: from an Aristotelian theory to a Newtonian one, for example. This is accomplished by placing students in situations where they are confronted by evidence which conflicts their personal theories, thus exposing the falsehoods of their intuitive assumptions. Others, such as diSessa (1988), propose that students do not have any "theory" in any sense of the word, but rather claim ownership to a collection of knowledge fragments, some of which are seen by experts as "correct" and "useful" and others as "incorrect" and "detrimental." In contrast, the instructional approach here is to capitalize and build on the student's knowledge base in order to build "a new and deeper systematicity" (diSessa, 1988, p. 51) where old knowledge is traded for new while a mental framework or model is constructed. In either case, problem-solving will be influenced by a student's mental model of the task or domain. Without an established and explicit mental model, students are likely to revert to misconceptions when solving novel problems, even though they may demonstrate mastery of specifically learned facts or procedures when tested.

Intrinsic Motivation

Constructivists make the assumption that learners must find an activity, within a microworld or otherwise, intrinsically motivating in order for the process of equilibration to occur. Lawler (1982) has suggested that microworlds, like those presented in LOGO, are successful because they produce "neat phenomena," or "phenomena that are inherently interesting to observe and interact with" (p. 141). However, constructivists offer little guidance on this issue to designers of microworlds. Turtle geometry, for example, may capture an innate human interest in the visual appeal of graphics, although this is totally speculative.

What motivates an individual to initiate and complete a task? A strict behaviorist would derive this empirically by trying a variety of reinforcements and observing which best sustains the desired response. School environments typically use a variety of reinforcements, such as praise, rewards, and grades. All of these are examples of extrinsic motivators, in that an activity is engaged in order to get the promised incentive, whether it be a star pasted to a school paper, or a good grade on a report card. We all need extrinsic motivators, such as a paycheck, at times in order to get an important job done which we may otherwise not choose to complete.

In contrast, people tend to complete other tasks without the promise of a reward. These tasks or activities have their own inherent appeal and people do not need coaxing to participate. Typical examples include leisure activities such as reading, playing cards, cooking, or crafts. In each of these cases, participation in the activity itself is reward enough. Such activities are intrinsically motivating in that the nature of the activity or domain itself produces significant reason to participate and external reinforcement or pressure to participate is not required (Kinzie & Sullivan, 1989; Maehr, 1976). Of course, an activity which is intrinsically motivating to one person may require extrinsic motivators for another.

Instruction which is intrinsically motivating relies on incentives which are student-centered, rather than external lesson reinforcement. In fact, the addition of extrinsic incentives, such as grades or other rewards, can often undermine or destroy the intrinsic appeal of many activities for students, and, in effect, turn "play into work" (Condry, 1977; Greene & Lepper, 1974; Lepper, Greene, & Nisbett, 1973). Malone (1981) has suggested a framework of intrinsically motivating instruction based on *challenge*, *curiosity*, and *fantasy*. Malone's discussion is particularly relevant here because he has specifically applied this framework to the design of computer games which often share characteristics of microworlds.

Challenge in an activity can usually be induced by establishing goals which have uncertain outcomes. The goal should be personally meaningful to the learner such that the skills or knowledge being learned are viewed as important because they help to achieve the goals. Tasks need to be designed to be optimally challenging, such that they are not too easy or too difficult. But perhaps most importantly, the tasks should elicit feelings of competence, or self-efficacy, in solving problems which students perceive as relevant and important. This enhances one's self-concept and leads to a feeling of control over one's own success (Weiner, 1979).

Closely related to challenge is curiosity. A person's curiosity is usually piqued when an activity is viewed as novel or moderately complex. In addition, curiosity is also usually increased by activities which offer a certain level of surprise. This occurs when the expected and actual outcome of an activity are different or incongruent, a phenomenon which Berlyne (1965) has termed "conceptual conflict." Again, however, both challenge and curiosity produced by a conceptual conflict must be optimally maintained to be effective. A task which is perceived as too easy quickly loses appeal, and a task which is seen as too demanding is avoided. Likewise, a conceptual conflict between expected and actual task outcomes can make a learner seek to resolve the conflict but can quickly lead to frustration if the conflict is too confusing or bewildering. Norman (1978) has termed optimal levels of conceptual conflict as "critical confusion."

Fantasy entails providing learners with a meaningful context for learning which is easy to augment with their imaginations. It is meaningful in the sense that it offers a very personal degree of fascination and intrigue which is easily transferred to play activities. Fantasy educes mental images of a context in which a learner is not actually present. The use of fantasy is common in theme parks and children's films and television programs. Malone (1981) has termed instruction which uses fantasies in this way as *intrinsic fantasies* because the skills to be learned are integral with the fantasy, such as learning about latitude and longitude in order to locate a "pirate's sunken treasure."

These characteristics of intrinsic motivation are all similar to Keller's (1983) synthesis of affective attributes which form the basis of his ARCS (Attention, Relevancy, Confidence, and Satisfaction) model of motivational instructional design. The degree to which these determinants are present in instruction and how they relate to differences in learning are not clear. Some examples of differences in learning given instruction which does or does not provide intrinsic incentives are the intensity of a learner's arousal and attention to a task, the depth of involvement, and task persistence (Lepper, 1985). Each of these in turn would influence the amount of time learners actually spend on-task and also the level at which mental processing occurs (Maehr, 1976; Malone, 1981; Rotter, 1954).

An Example of Merging Constructivism and Direct Instruction

Although many of the principles of constructivism offer promise in the development of successful learning environments, practical applications are difficult to draw out given the typical constraints of most school and training situations. The purpose of this final section is to synthesize the principles discussed so far by presenting a working prototype of an instructional development project based on the merger of constructivism and direct instruction. This project was founded on compromises between the extreme views of both camps such as accepting the constructivist philosophy that learners should be given rich and powerful environments to build and transform mental structures while maintaining the practicality of current educational conditions which the instructivist position supports. Table 1 summarizes a series of design considerations which are the foundation of this synthesis.

The project involved the development of a computer software package called *Space Shuttle Commander* (SSC). SSC is a "mini-course" which teaches Newton's laws of motion to elementary and middle school students to achieve the learning objectives listed in Table 2. SSC combines the use of computer tutorials and microworlds. The microworlds comprise a series of simulations/games called "missions" in which students pilot an animated "shuttle," as represented in Figure 1. The tutorials consist of a series of "flight lessons" which formally introduce the laws of motion. SSC takes a nonmathematical approach and instead concentrates on concept formation. Portions of both the tutorials and microworld activities in SSC have been validated through a series of ongoing studies involving the role of computer graphics in learning from animated lesson presentations and visually-based simulations (Rieber, 1990a; Rieber, in press; Rieber, Boyce, & Alkindi, in press).

SSC is a direct application of a physics microworld designed by Andy diSessa (1982) using LOGO which involves a dynamic turtle, or "dynaturtle." The dynaturtle is familiar to the turtle described earlier, except that it has one additional characteristic — velocity. The dynaturtle acts as a free-floating Newtonian particle where students can explore motion principles in a simulated frictionless, gravity-free environment. Examples of other projects based on the dynaturtle microworld include White (1984) and White and Horwitz (1987). SSC is distributed through the fifteen NASA Teacher Resource Centers located nationwide (Rieber, 1991).

Table 1

Some Considerations in the Design of Computer-Based Microworlds to Support Direct Instruction

- Provide a meaningful learning context which supports intrinsicially motivating and self-regulated learning.
- Establish a pattern where the learner goes from the "known to the unknown."
- Provide a balance between deductive and inductive learning.
- Emphasize the usefulness of errors.
- Anticipate and nurture incidental learning.

SSC attempts to take advantage of the strengths of tutorials and simulations (microworlds) while minimizing their weaknesses by combining them in a cohesive, yet flexible, course. For example, traditional tutorials, combined with drill and practice, while good at presenting content matter in a logical order and providing practice, are usually monotonous and unimaginative and usually only promote learning at surface levels. Microworlds, presented in the context of simulations and games, offer the potential for intrinsically motivating and cognitively rewarding activities, however, they make the presentation and validation of a predetermined set of learning objectives difficult.

Each of the design considerations in Table 1 will be discussed in the context of SSC.

Provide a meaningful learning context which supports intrinsically motivating and self-regulated learning. The success of any microworld depends on a context in which "neat phenomena" are produced. As the motivation literature suggests, the design of activities in which learners demonstrate commitment and perseverance in the thoughtful completion of a task depends on the degree to which the activity is

Table 2

Lesson objectives for Space Shuttle Commander (SSC)

Terminal Objectives. At the end of SSC, each student will be able to:

- list Newton's three laws of motions;
- apply Newton's three laws of motions given a variety of gravity-free/frictionless situations in one- and two-dimensional space using questioning and simulation techniques;
- state examples and nonexamples of forces;
- describe and explain verbally the functions and control of the simulated space shuttle using vocabulary and concepts from Newtonian mechanics.
- choose to return to the various shuttle missions given the freedom of choice.

Enabling Objectives. The following enabling objectives describe the expected learning outcomes for the flight lessons and the first four missions.

At the end of Flight Lesson 1 and Mission 1, each student will be able to:

- state and apply Newton's first law given the context of no gravity and no friction (an object in motion will remain in motion and an object at rest will remain at rest).

At the end of Flight Lesson 2 and Mission 2, the student will be able to:

- apply an application of Newton's second law where, in one-dimensional space, when an object at rest is put into motion by a force, an equal force must applied in the opposite direction to stop the object, given the context of no gravity and no friction.

At the end of Flight Lesson 3 and Mission 3, each student will be able to:

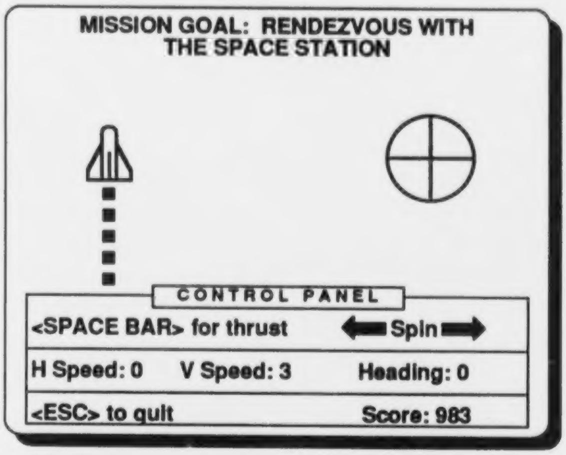
- apply an application of Newton's second law where, in one-dimensional space, when unequal forces act on an object in opposite directions, the final speed and trajectory of the object is the result of the difference between the sum of the opposing forces, given the context of no gravity and no friction;
- apply an application of Newton's second law where, in one-dimensional space, the initial acceleration of an object depends on the acting force and the mass of the object, given the context of no gravity and no friction.

At the end of Flight Lesson 4 and Mission 4, each student will be able to:

- apply the principle that, in two-dimensional space, when forces acting on an object occur at right angles to one another they act independently on the object, given the context of no gravity and no friction.

At the end of Flight Lesson 5, each student will be able to:

- state Newton's third law and give at least one example of its application.



MISSION GOAL: RENDEZVOUS WITH THE SPACE STATION

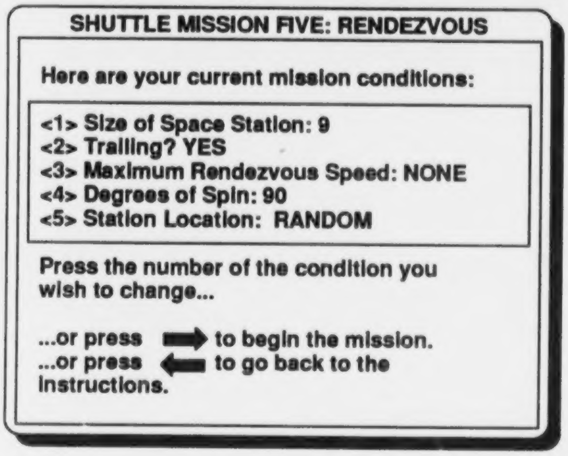
CONTROL PANEL

<SPACE BAR> for thrust ⬅ Spin ➡

H Speed: 0 V Speed: 3 Heading: 0

<ESC> to quit Score: 983

FIGURE 1. *A representation of the computer screen during an episode of "Mission 5: Rendezvous." The animated "shuttle" is under student control. Arrow keys rotate the shuttle in 90 degree increments and the space bar gives the shuttle a "kick" or thrust in the direction it is pointing. The goal of this mission is to manuever the shuttle to the space station.*



SHUTTLE MISSION FIVE: RENDEZVOUS

Here are your current mission conditions:

<1> Size of Space Station: 9
<2> Trailing? YES
<3> Maximum Rendezvous Speed: NONE
<4> Degrees of Spin: 90
<5> Station Location: RANDOM

Press the number of the condition you wish to change...

...or press ➡ to begin the mission.
...or press ⬅ to go back to the instructions.

FIGURE 2. *Sample screen showing the various mission "conditions" of Rendezvous which can be changed to increase the difficulty and challenge of the lesson. (Not drawn to scale.)*

perceived as relevant and its completion as personally satisfying (Keller & Suzuki, 1988). By definition, a meaningful learning context is an intensely personal affair. The goal in education, however, is to discover contexts which have a wide appeal to learners of varying interests and aptitudes. LOGO, for example, seems to attract the attention of children through the use of interactive computer graphics to produce interesting visual designs.

My experience has shown that imagining to be an astronaut appears to be a common fantasy in which children and adults often indulge. The successes and failures of the world's space flights during the past 30 years offer an intriguing context for people's imaginations to go on fantastic voyages in outer space. Beyond all of the fantasy is the very real practicality and necessity of working in space which is continually documented in the media. SSC tries to capitalize on all this by encouraging students to fantasize about being the commander of the space shuttle. The purpose of the "missions" is to induce and encourage the fantasy in a series of structured activities, each building on the other. The flight lessons are meant to represent "flight school." A similar example of using a meaningful context for learning is the *Voyage of the Mimi*, a math and science curriculum set in the context of whale exploration.

An important conclusion of Malone's (1981) research is that students need to be provided with the ability to continually increase the challenge of an activity in order to maintain intrinsic appeal. Missions 5 and 6, for example, provide students with the opportunity to vary the "mission conditions" (such as target size, target location, shuttle rotation, and other game constraints) in order to increase the difficulty of the activity, such as shown in Figure 2. For example, it is much simpler to control the shuttle in a mission occurring in two-dimensional space when the shuttle's rotation is constrained to 90 degree increments. Control of the shuttle becomes considerably more difficult and complex when control is changed to 45 or 30 degree increments. In addition, scorekeeping features are provided to allow students to chart their progress under any given set of conditions. Scorekeeping is another feature noted by Malone (1981) as increasing intrinsic appeal in computer games.

Establish a pattern where the learner goes from the "known to the unknown." Meaningfulness can also be interpreted as the degree to which students can link new ideas to prior knowledge, or what the student already knows. The degree to which information is related to prior knowledge is among the most important determinants of learning (Ausubel, 1968). Bruner (1966), for example, suggests a "spiral" approach to learning where the simplest and most general ideas are introduced first to learners in highly interactive and concrete ways. These ideas are then successively reintroduced to students to provide increasing levels of detail. This is similar to Ausubel's (1968) *progressive differentiation*, and Reigeluth's elaboration theory (Reigeluth & Stein, 1983), where general ideas grasped early by learners help them to subsume more detailed ideas introduced later.

SSC can act as an important bridge between experiential learning with a dynaturtle and formal physics (as usually taught in schools). Paradoxically, traditional physics instruction based on highly mathematical models usually removes the influences of friction and gravity in order to "simplify" learning for students. While mathematical formulas may be simplified using this convenience, students' conceptual understanding is usually muddled since the removal of friction and gravity is totally outside of their daily experiences. SSC tries to provide students with a sound conceptual understanding of Newtonian principles to act as "anchoring posts" for subsequent instruction.

In addition, the simulated space shuttle, like its cousin the dynaturtle, acts as a transitional object between the learner and Newtonian physics allowing the exploration of physics ideas within a context that is relevant and fantasy-driven as well as body syntonic.

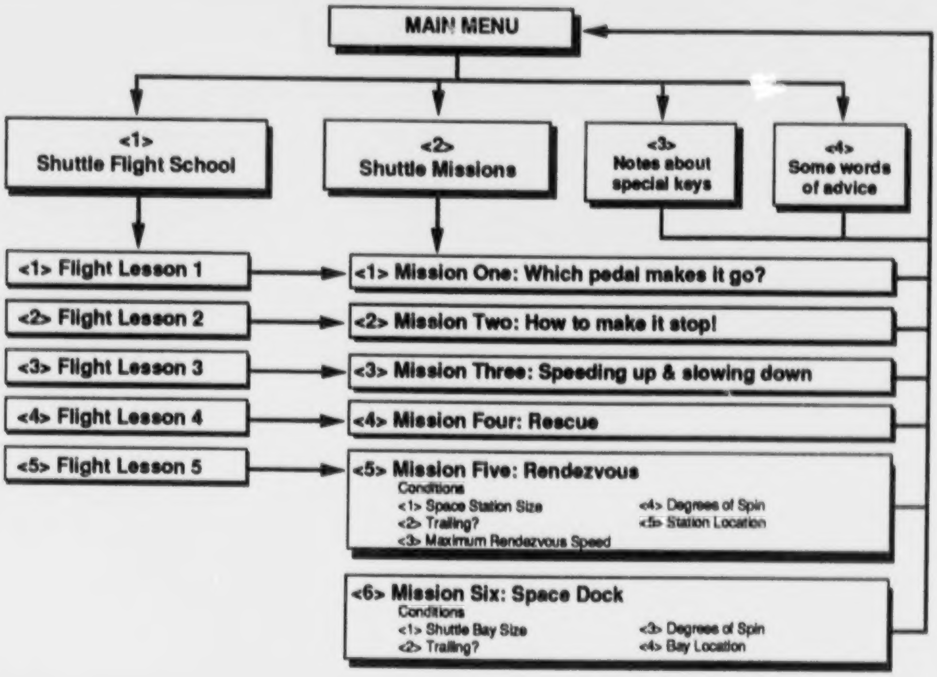
Provide a balance between deductive and inductive learning. Historically, ISD approaches are based on deductive learning models where the instructional goal is to identify as precisely as possible a set of learning objectives and bring the learner to mastery of those objectives as directly and efficiently as possible (Reigeluth, 1983). Students are given a set of "truths" to learn and apply. For example, a rule, such as Newton's first law is explicitly presented with a variety of supportive examples and activities. Students then practice using and applying the rule, and eventually are tested on their understanding. If mastery is achieved, students proceed to the next objective in the hierarchy; if mastery is not achieved, the students are either provided with remediation on the objective or are exited to receive instruction on prerequisites. In contrast, a constructivist approach involves largely inductive strategies where instances of a rule are provided with the goal that the learner will induce the rule for themselves. By so doing, learners make each "truth" they discover their own. The supposition is that this leads to deeper levels of understanding and also intrinsically motivates students to persist in the task.

The problems with extreme interpretations of either approach are obvious. Strict deductive approaches can lead to instructional designs which are unnecessarily self-limiting and which are inclined to assign a passive role to the learner. Instructional designs are likely to resemble one another resulting in activities which lack imagination and innovation. Deductive approaches are much easier to apply when learning outcomes involve verbal information because there is little need for interpretation or inference on the part of the learner. Conversely, strict inductive approaches resemble a "sink or swim" philosophy where learners may become either bored or frustrated if they are unable to generalize from the instances provided. In addition, novices often need structure or guidance which purely inductive experiences do not provide, a problem frequently encountered by designers of hypertext (Jonassen, 1986; Tripp & Roby, 1990). Inductive activities also require an attitude of playfulness and exploration which older children or adults may find as risky or threatening when being judged by peers.

In SSC, the activities are designed to follow either a deductive or inductive approach, as well as allowing the provision to switch between each approach. On the surface SSC is laid out essentially in a deductive fashion. The tutorials or "flight lessons" are hierarchically organized where later skills build on earlier ones, as Table 2 and the lesson flowchart in Figure 3 show. Each flight lesson "teaches" the respective objectives according to conventional instructional design and each "mission" acts as a suitable practice activity for each lesson (Gagné, Briggs, & Wager, 1988). Therefore, students can go through SSC in the classic deductive fashion, starting with the first Flight Lesson.

However, each mission acts as a stand-alone microworld which simulates one or more particular aspects of Newton's laws of motion as suggested by the lesson objectives shown in Table 2. For example, Missions 1, 2, and 3 take the learner through a series of activities which introduce the simplest aspects of Newton's first and second laws. At first, structure is heavily imposed, but then reduced as the learner establishes a foothold with these concepts. The first three

FIGURE 3. *A lesson flowchart of* Space Shuttle Commander.


MAIN MENU
<1> Shuttle Flight School
<2> Shuttle Missions
<3> Notes about special keys
<4> Some words of advice
<1> Flight Lesson 1
<2> Flight Lesson 2
<3> Flight Lesson 3
<4> Flight Lesson 4
<5> Flight Lesson 5
<1> Mission One: Which pedal makes it go?
<2> Mission Two: How to make it stop!
<3> Mission Three: Speeding up & slowing down
<4> Mission Four: Rescue
<5> Mission Five: Rendezvous
Conditions
<1> Space Station Size
<2> Trailing?
<3> Maximum Rendezvous Speed
<4> Degrees of Spin
<5> Station Location
<6> Mission Six: Space Dock
Conditions
<1> Shuttle Bay Size
<2> Trailing?
<3> Degrees of Spin
<4> Bay Location

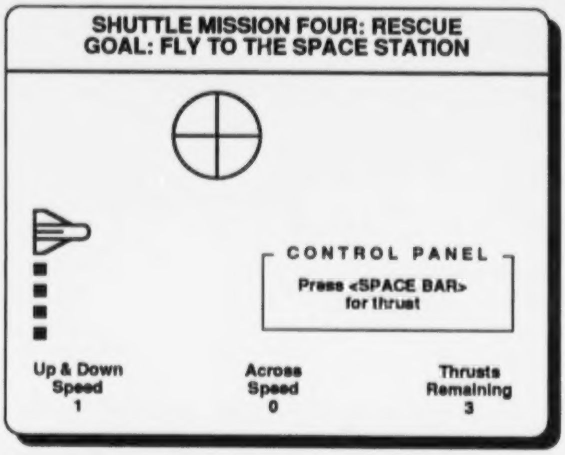
missions further constrain the learner's experience to one dimension. Mission 4, as illustrated in Figure 4, introduces the effects of two dimensions in a highly structured way so as to make the simplest relationships of two-dimensional motion as explicit as possible. The artificial restraints on this microworld are presented in a "rescue" context where students are asked to imagine that their shuttle has been disabled by a collision with a meteorite. The collision has resulted in limited control of the shuttle.

It is possible, therefore, to have students begin to understand Newtonian mechanics by only having them explore the missions in SSC. The flight lessons become resources to provide formal accounts of the information or principles which students actually use on the missions. Flight lessons would be consulted only when clarification is needed, either the result of curiosity or confusion.

The most radical example of using an inductive approach with SSC would place students in one of the last two missions, such as "Rendezvous," depicted in Figure 1, or "Space Dock", depicted in Figure 5. Both are fully detailed simulations. Many students quickly become successful "commanders" of the space shuttle in these missions and usually seek to increase the challenge of the activities by changing the mission conditions. Students who are not successful would be encouraged by on-line coaching (and hopefully by the teacher as well) to go through earlier missions or to consult the flight lessons.

In practice, most students choose a combination of deductive and inductive strategies. The teacher plays an important role to facilitate the best path for each student. A teacher who already has developed an established unit on Newton's laws might use SSC to supplement or supplant part of their curriculum. Other teachers might use SSC to fully guide students through the unit, but should feel free to explore other activities and ideas as well. SSC, like the original dynaturtle microworld, is designed to bring students in contact with ideas which conventional educational wisdom might suggest they are not ready for.

Emphasize the usefulness of errors. Errors are an inevitable part of the learning process without which learning would not occur. Inductive learning theories promote the ability of learners to detect errors and then incorporate the information learned in subsequent trials. Using errors in inductive learning is not meant to be a haphazard process, but is usually quite systematic. When confronted with the task of learning about two or more unknown variables, a learner must be able to isolate and manipulate each of the variables while

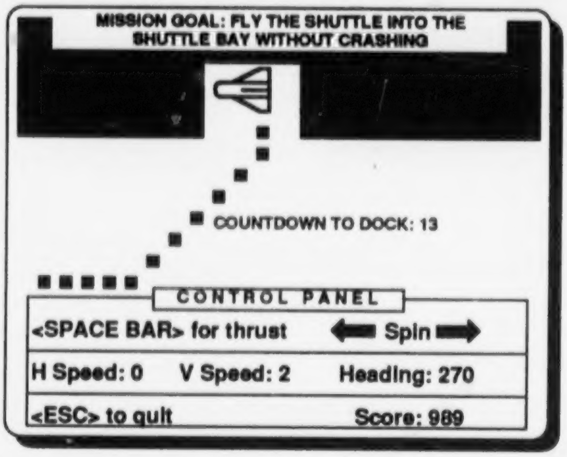


SHUTTLE MISSION FOUR: RESCUE
GOAL: FLY TO THE SPACE STATION

CONTROL PANEL
Press <SPACE BAR> for thrust

Up & Down Speed 1
Across Speed 0
Thrusts Remaining 3

FIGURE 4. *A representation of the computer screen during "Mission 4: Rescue." The goal of the mission is to manuever a "disabled" shuttle to the space station. Students do not have control over the shuttle's rotation, and they only have enough fuel for three bursts of thrust. These constraints make the relationship of orthogonal forces more apparent. (Not drawn to scale.)*



MISSION GOAL: FLY THE SHUTTLE INTO THE SHUTTLE BAY WITHOUT CRASHING

COUNTDOWN TO DOCK: 13

CONTROL PANEL
<SPACE BAR> for thrust ⬅ Spin ➡
H Speed: 0 V Speed: 2 Heading: 270
<ESC> to quit Score: 989

FIGURE 5. *A representation of the computer screen during "Mission 5: Space Dock." The goal of this mission is to manuever the shuttle into the shuttle "bay" of the space station without touching the walls of the bay or the station. (Not drawn to scale.)*

holding all other variables constant. This systematic process of error handling forms the basis of hypothesis forming and testing, as shown in the research on concept formation (see Mayer, 1983, for a review). The ability of a learner to focus on one variable at a time while understanding the relationship among two or more variables is an example of decentered thinking (as compared to egocentric thinking), of which the ability to conserve is an example (Slavin, 1988).

Unfortunately, many learning tasks contain so many individual variables that a novice learner would soon be inundated with information, leading to frustration. Microworlds offer a way to structure a learning experience so that a finite set of variables are introduced at any one time. When these variables are mastered, additional variables can be introduced. Piaget called this *variable-stepping* which is an essential characteristic of forming and testing hypotheses. In classical mechanics, for example, problems can be simplified when presented in one rather than two dimensions. In a computer microworld, like the SSC missions, this is accomplished in a number of ways. Certain computer commands can be activated or deactivated, depending on which variables are to be explored in a particular activity. Other commands can be constrained so as to produce a particular effect, such as limiting the rotation of the shuttle to 180 degree increments so that one-dimensional motion is produced.

Errors cannot be useful unless the goal of an activity is clearly known. If the goal is ambiguous, then all available feedback will also be ambiguous. In LOGO, for example, students often work towards the completion of a graphic which they have chosen, such as a house or a car. The graphical feedback they receive from the turtle is continually judged against their individually defined goals. In the best of cases, goal monitoring is automatic and intrinsic. The best rule of thumb for microworld design is to provide the simplest and clearest goals. An example of a mission goal in

SSC is "fly the shuttle to the space station."

Finally, a variety of secondary feedback features can also be used to complement primary feedback mechanisms. For example, the shuttle in SSC automatically leaves a trail in order to show the history of its travel path (although this feature can be turned off as many people find the trailing distracting at times). Verbal feedback can also be presented in tandem with visual feedback. SSC provides a "control panel" showing such information as the shuttle's vertical and horizontal speed and the shuttle's heading. Verbal feedback can be important when novices have difficulty seeing slight visual changes, such as when the shuttle is moving very slowly.

The importance of an attitude change toward errors on the part of teachers and students must be underscored. The negative connotation of errors is firmly rooted in school environments, and simply stating that errors should be considered useful will not change this. The attitude toward errors in learning is probably one of the starkest differences between behavioral and cognitive learning perspectives.

Anticipate and nurture incidental learning. Constructivist approaches accept the fact that learning does not necessarily flow from a fixed sequence of ideas. In LOGO programming, for example, mistakes or "bugs" often lead to other interesting phenomena and students often choose to abandon an original programming project in lieu of projects which follow unexpected results. However, this can make it difficult to identify and document mastery of learning goals. In contrast, instructivist approaches strive to take a group of learners through a sequence of predetermined learning objectives to the point of ignoring any learning which may be incidental to these objectives. Learners are not only less likely to explore a wider array of learning experiences, but are actually discouraged from doing so. Carefully designed microworlds should expect and encourage incidental learning to occur within design parameters. The teacher's role is again very important here. A teacher can help to channel incidental learning to serve the lesson or unit's terminal goals by helping to illustrate their relevancy. Likewise, students are apt to get sidetracked by incidental paths and a teacher should redirect a student back to relevancy if the incidental learning is too esoteric or counterproductive.

One example of the two edges of the incidental learning "sword" is provided by Rieber (in press) in which incidental learning from the tutorial sections of SSC was studied. Students in the study were given a tutorial on a simple application of Newton's second law where the acceleration of an object with a constant mass varies depending on the force that is applied to it. However, students given an animated account of this principle also were taught another application of this physics law incidentally. Through animation, students saw the motion consequences when the same size force is applied to objects of different mass. Fourth grade students successfully extracted this incidental application even though no formal attempt was made to teach them the rule. However, these students continued to apply this incidental information to other *inappropriate* contexts, such as gravitation problems. The conclusion here is that students who learn in incidental ways apply this information in a variety of contexts: some of which may be appropriate and constructive to a larger set of learning goals; some of which may actually undermine some learning goals, such as by promoting misconceptions.

Closing

This paper has discussed the educational dilemma of promoting a constructivist philosophy within an educational system based on direct instructional methods. This is not necessarily the paradox it appears to be. Rather than being a destructive influence on instructional technology, it is contended that this tension is healthy and useful as well as necessary in the maturation of the field. This tension acts as a "dithering device" which motivates reflectivity and growth and helps to eliminate stagnation and professional nepotism. The ability to learn and otherwise progress in cognitive ways is a natural, innate and *personal* process for people and one which the constructivist approach advocates. Yet, most school instruction is often centered solely around the content and not the student. The instructivist approach sees learning as the mastery of a series of objectives where the mastery of one objective is seen as the starting point for the next. This approach risks the danger of focusing on the content to be learned, instead of the learner and the learning experience. Although the goals of education in terms of content are quite similar in each approach, the means to achieve those goals are radically different. It is suggested that computer microworlds offer an interesting compromise between the instructivist and constructivist approaches. Microworlds can be designed in such a way as to give users exploratory experiences within a carefully controlled range of concepts and principles and thereby offer a practical compromise between instructivism and constructivism.

References

Abelson, H., & diSessa, A. (1981). *Turtle geometry: The computer as a medium for exploring mathematics.* Cambridge, MA: MIT Press.

Ausubel, D. (1968). *Educational psychology: A cognitive view.* New York: Holt, Rinehart & Winston.

Berlyne, D. (1965). *Structure and direction in thinking.* New York: John Wiley.

Case, R., & Bereiter, C. (1984). From behaviourism to cognitive behaviourism to cognitive development: Steps in the evolution of instructional design. *Instructional Science, 13*, 141-158.

Champagne, A., Klopfer, L., & Anderson, J. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics, 48*, 1074-1079.

Chi, M., Feltovich, P., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*, 121-152.

Clark, R. (1980). Do students enjoy the instructional method from which they learn the least? Antagonism between enjoyment and achievement in ATI studies. (ERIC Document Reproduction Service No. ED 188 601)

Clements, D. (1985). Research on LOGO in education: Is the turtle slow but steady, or not even in the race? *Computers in the Schools, Summer/Fall*, 55-71.

Condry, J. (1977). Enemies of exploration: Self-initiated versus other-initiated learning. *Journal of Personality and Social Psychology, 35*, 459-477.

Dede, C. (1987). Empowering environments, hypermedia and microworlds. *The Computing Teacher, 15*(3), 20-24, 61.

diSessa, A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science, 6*, 37-75.

diSessa, A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, NJ: Erlbaum.

diSessa, A., & Abelson, H. (1986). Boxer: A reconstructible computational medium. *Communications of the ACM*, 29(9), 859-868.

Eylon, B., & Linn, M. (1988). Learning and instruction: An examination of four research perspectives in science education. *Review of Educational Research, 58*, 251-301.

Forman, G., & Pufall, P. (Eds.). (1988a). *Constructivism in the computer age.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Forman, G. & Pufall, P. (1988b). Constructivism in the computer age: A reconstructive epilogue. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 235-250). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gagné, R. (1985). *The conditions of learning* (4th ed.). New York: Holt, Rinehart, & Winston.

Gagné, R. & Dick, W. (1983). Instructional psychology. *Annual Review of Psychology, 34*, 261-295.

Gagné R., & Glaser, R. (1987). Foundations in learning research. In R. Gagné (Ed.), *Instructional Technology: Foundations* (pp. 49-83). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gagné, R. & Merrill, M. (1988, January). *Presidential session: A conversation on instructional design.* Presented at the annual convention of the Association for Educational Communications and Technology, Dallas, TX.

Gentner, D., & Stevens, A. (Eds.). (1983). *Mental models.* Hillsdale, NJ: Lawrence Erlbaum.

Greene, D., & Lepper, M. (1974). How to turn play into work. *Psychology Today, 8*, 49-54.

Hannafin, M., & Rieber, L. (1989). Psychological foundations of instructional design for emerging computer-based instructional technologies: Part I. *Educational Technology Research & Development, 37*(2), 91-101.

Howe, J., O'Shea, T., & Plane, F. (1980). Teaching mathematics through LOGO programming: An evaluation study. In R. Lewis & E. Tagg (Eds.), *Computer assisted learning: Scope, progress, and limits* (pp. 85-102). Amsterdam, Holland: North-Holland.

Jonassen, D. (1986). Hypertext principles for text and courseware design. *Educational Psychologist, 21*(4), 269-292.

Keller, J. (1983). Motivational design of instruction. In C.M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current states* (pp. 383-434). Hillsdale, NJ: Lawrence Erlbaum Associates.

Keller, J., & Suzuki, K. (1988). Use of the ARCS motivation model in courseware design. In D. Jonassen (Ed.), *Instructional designs for microcomputer courseware* (pp. 401-434). Hillsdale, NJ: Erlbaum.

Kinzie, M., & Sullivan, H. (1989). Continuing motivation, learning control, and CAI. *Educational Technology Research & Development, 37*(2), 5-14.

Knirk, F., & Gustafson, K. (1986). *Instructional technology: A systematic approach to education.* New York: Holt, Rinehart, & Winston.

Kobayashi, S. (1986). Theoretical issues concerning superiority of pictures over words and sentences in memory. *Perceptual and Motor Skills, 63*, 783-792.

Kulhavy, R. (1977). Feedback in written instruction. *Review of Educational Research, 47*, 211-232.

Lane, D. (1980). Incidental learning and the development of selective attention. *Psychological Review, 87*, 316-319.

Lawler, R. (1982). Designing computer-based microworlds. *Byte, 7*(8), 138-160.

Lepper, M. (1985). Microcomputers in education: Motivational and social issues. *American Psychologist, 40*, 1-18.

Lepper, M., Greene, D., & Nisbett, R. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the overjustification hypothesis. *Journal of Personality and Social Psychology, 28*, 129-137.

Lockard, J., Abrams, P., & Many, W. (1990). *Microcomputers for educators* (2nd ed.). Glenview, IL: Scott, Foresman/Little, Brown HIgh Education.

Lukas, G., & Lukas, J. (1986). *LOGO: Principles, programming, and projects.* Monterey, CA: Brooks/Cole Publishing Company.

Maehr, M. (1976). Continuing motivation: An analysis of a seldom considered educational outcome. *Review of Educational Research, 46*, 443-462.

Malone, T. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 4*, 333-369.

Mayer, R. (1983). *Thinking, problem solving, cognition.* New York: W. H. Freeman & Company.

Mayer, R. (1984). Aids to text comprehension. *Educational Psychologist, 19*(1), 30-42.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.

Norman, D. (1978). Notes toward a theory of complex learning. In A. Lesgold, J. Pellegrino, S. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction* (pp. 39-48). New York: Plenum Press.

Norman, D. (1982). *Learning and memory.* San Francisco: W.H. Freeman & Company.

Norman, D. (1983). Some observations on mental models. In D. Gentner, & A. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, NJ: Lawrence Erlbaum.

Norman, D. (1988). *The psychology of everyday things.* New York: Basic Books.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: Basic Books.

Papert, S. (1981). Computer-based microworlds as incubators for powerful ideas. In R. Taylor (Ed.), *The computer in the school: Tutor, tool, tutee* (pp. 203-210). New York, NY: Teacher's College Press.

Papert, S. (1987). Computer criticism vs. technocentric thinking. *Educational Researcher, 16*(1), 22-30.

Papert, S. (1988). The conservation of Piaget: The computer as grist to the constructivist mill. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 3-13). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. (Ed.). (1983). *Instructional-design theories and models: An overview of their current status.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Reiser, R. (1987). Instructional technology: A history. In R. Gagné (Ed.), *Instructional Technology: Foundations* (pp. 11-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rieber, L. (in press). Computer animation, incidental learning, and continuing motivation. *Journal of Educational Psychology.*

Rieber, L. (1990a). Animation in computer-based instruction. *Educational Technology Research & Development, 38*(1), 77-86.

Rieber, L. (1990b). Using computer animated graphics in science instruction with children. *Journal of Educational Psychology, 82*, 135-140.

Rieber, L. (1991). *Space shuttle commander* [Computer Software]. Washington, DC: NASA Educational Technology Branch.

Rieber, L., Boyce, M., & Alkindi, M. (in press). The effects of computer-based interactive visuals as orienting and practicing activities on retrieval tasks in science. *International Journal of Instructional Media.*

Rogers, C. (1969). *Freedom to learn.* Columbus, OH: Merrill.

Rotter, J. (1954). *Social learning and clinical psychology.* Englewood Cliffs, NJ: Prentice-Hall.

Skinner, B. (1968). *Technology of teaching.* New York: Meredith Publishing Company.

Slavin, R. (1988). *Educational psychology: Theory into practice.* Englewood Cliffs, NJ: Prentice-Hall.

Steinberg, E. (1977). Review of student control in computer-assisted instruction. *Journal of Computer-Based Instruction, 3*, 84-90.

Suppes, P., & Morningstar, M. (1972). *Computer-assisted instruction at Stanford, 1966-1968: Data, models, and evaluation of the arithmetic programs.* New York: Academic Press.

Taylor, R. (Ed.). (1980) *The computer in the school : tutor, tool, tutee.* New York, NY: Teacher's College Press.

Tripp, S., & Roby, W. (1990). Orientation and disorientation in a hypertext lexicon. *Journal of Computer-Based Instruction, 17*, 120-124.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*, 3-25.

White, B. (1984). Designing computer games to help physics students understand Newton's laws of motion. *Cognition and Instruction, 1*, 69-108.

White, B., & Horwitz, P. (1987). *ThinkerTools: Enabling children to understand physical laws* (Report No. 6470). Cambridge, MA: BBN Laboratories.

Title:

Adjunct Questions and Rule Learning

Author:

Steven Rodriguez

Educators tend to agree about the criticality of integrating questions in learning material. Questions provide learners with opportunities to apply or process the content of instruction, thereby supporting learning. Early research demonstrated that adjunct questions do have positive influences on acquisition and retention of factual information (Hamilton, 1986; Rickards & Denner, 1978; Watts & Anderson, 1971).

Recent authors--looking beyond verbal information learning--have expressed interest in the potential of adjunct questions to facilitate learning in the intellectual skills domain (cf. Andre, Mueller, Womack, Smid, & Tuttle, 1980; Jonassen, 1985; Carrier & Fautsch-Patridge, 1981). The present discussion considers this research area and offers the following: 1) a rationale for pursuing research regarding the effects of adjunct questions on rule learning; 2) a framework for conceptualizing such research; 3) a description of a related research study; and 4) general recommendations for future research.

Adjunct Questions and Intellectual Skills: Why Bother?

The usefulness of continuing to investigate how questions influence rote recall has been challenged. As early as 1971, Watts and Anderson noted, "Whether or not a student is able to accurately recognize... strings of words is educationally trivial" (p. 387). Others extend this argument, stating that the practical educator should be more concerned with whether students understand concepts and principles than whether they can recall factual details (Anderson & Biddle, 1975). These views are indicative of the interest that presently exists concerning use of adjunct questions to support learning in the intellectual skills domain as described by R. Gagne (1985). Intellectual skills include concept learning, rule using, and problem solving.

Existing research provides a second more specific reason for pursuing research in this area. Minimal consensus has emerged from the research that has addressed the impact of adjunct questions on intellectual skill learning. Studies concerned with concept learning illustrate this point. Some studies have indicated that adjunct questions questions facilitate concept learning (cf., Watts & Anderson, 1971; E. Gagne, Broughton, Eggleston, Holmes, Hawkins, & Sheldon,1979). Others have not.

In their seminal study, Watts and Anderson (1971) had subjects answer one of three types of questions. These questions entailed either identifying a previously presented example of a principle described in the text, identifying a new example of a principle, or identifying the name of the person associated with a given principle. Results indicated that application questions requiring identification of new examples facilitated subjects' ability to identify previously unencountered examples of the presented concepts. The researchers refer to the idea of *levels of processing* in suggesting that "... subjects answering application questions... were forced to engage in more thorough processing of the passages" (p. 392). They further note that application-level questions probably induced subjects to review passages in a manner that produced the deeper processing.

A study conducted by E. Gagne et al. (1979) also demonstrated the positive effects that higher-order questions may have on learning. In this study, subjects read passages on the concepts of autism and trochaic meter. Following each passage, subjects answered one of four types of questions requiring them to either summarize the passage, generate a new example of the concept, list the critical attributes of the concept, or identify a new example of the concept. Results support the view that adjunct questions requiring subjects to identify new examples facilitated successful classification of instances of the given concepts. The authors suggest that the questions accompanying the text included "close-in nonexamples" (p. 305) which may have caused subjects to carefully discriminate concept attributes.

Other authors dispute the view that adjunct questions promote concept learning. Andre et al. (1980) investigated whether including application questions in instruction would increase subjects' ability to apply the concepts being learned. Results of these and other studies (cf. Bing, 1982; Hamilton, 1986) appear to contradict the intuitively appealing idea that adjunct questions requiring subjects to apply content will result in superior criterion performance over either recall questions or no questions. While Andre and associates acknowledge the assumptions educators make about the perceived importance of asking higher-order questions, they conclude that "...there is little evidence that later application is influenced by the type of adjunct question" (p. 542). Hamilton offers a similar conclusion.

The confusion regarding the effects of adjunct questions on concept learning is indicative of the need for further research. Greater understanding of what kinds of questions best support intellectual skill learning will promote effective instructional design while adding to our knowledge of the learning process.

Active Role of the Learner

An overarching framework for conceptualizing research in the area of interest is presented in Figure 1. A major assumption inherent in the framework is this: the successful learner takes an *active* role in the learning process. Consistent with a general cognitive orientation toward research, this view is further reflected in the constructivist and generative conceptions of learning that some have offered.

---

insert Figure 1 here

The constructivist view holds that the learner builds new understanding and knowledge on the basis of existing knowledge (Pea, 1988). Motivational factors such as the perceived relevance of the instructional content (see Keller, 1983) also affect the degree of effort put forth by the learner in processing the presented information. This outlook is consistent with the so-called generative view of learning.

In a discussion of generative learning, Wittrock (1984) suggests that "... learners construct their own meaning from teaching" (p. 124). Thus learning with understanding requires that one determine the structure of what is to be learned and then generate relationships among the components. Further, learning is enhanced "...by constructing relationships between something familiar and something to be remembered" (p. 123).

Assimilation theory as described by Mayer (1980) is also consistent with the constructive and generative views. According to this theory, the key aspects of learning include receiving the to-be-learned material and activating an assimilative context grounded in meaningful, related past experience. All three perspectives--the constructivist, generative, and assimilative--stress that students must actively relate new information to existing knowledge if learning is to occur. This is sometimes referred to as the process of *elaboration* (Mayer, 1984; Jonassen, 1985).

At a conceptual level, it is useful to distinguish between the generative and mathemagenic views of learning as discussed by Jonassen (1985). The mathemagenic hypothesis essentially attributes meaningful learning to "...certain design attributes..." (p. 9) of the instruction such as adjunct questions that are conceived to induce the cognitive processes that are requisite to learning. Jonassen notes that this view "...assumes that learners are better led to learning than allowed to determine the relevance or meaning of material themselves" (p. 10). Conversely, the generative view asserts that learners confronting new information in instruction "... construct and assign meaning to that information based upon prior learning" (p.11).

Reigeluth (1983) notes the possibility of reconciling these two views via inclusion of cognitive strategy activators within instruction. These may be direct, embedded activators such as diagrams or analogies, or they may be or less direct, detached activators that advise the student to apply a given cognitive strategy (e.g., "Draw a diagram showing the relationship explained in the previous section"). Embedded activators are consistent with the mathemagenic conception of learning, detached activators with the generative view.

the layout, organization, or motivational appeal of the material may also influence learning;

2) **active role of learner:** the student's role may no longer be equated with the dash in the stimulus-response (S-R) or behaviorist conception of learning; the learner must actively process new information, elaborating and employing various cognitive strategies in light of expectations if meaningful learning is to occur;

3) **nature of the learning outcome:** future research may be most useful when it considers how various designs of adjunct questions influence learning of intellectual skills rather than rote fact learning: the functional usefulness of learning provides one possible indicator of the so-called ecological validity of research.

Description of Related Research

The present author is currently planning a research study that will be conducted in the spring of 1991. The purpose of the study is to ascertain the differential effects on learning of three types of adjunct questions incorporated in a text lesson on basic statistical concepts and rules. Treatments are derived from three separate theoretic perspectives: assimilation theory, forward transfer theory, and cognitive strategy activation.

Assimilative Context Activation

The *context activation* version of the instruction will include questions designed to require activation of an assimilative context as discussed by Mayer (1980). These are questions requiring that subjects relate the subject content to their own previous experience in providing a short written answer.

Assimilation theory entails three key processes conceived to support learning. First is the idea of reception: the learner must focus attention on information presented in the text in order for it to enter short term memory. Second is the idea of availability: the learner must have relevant existing knowledge that can serve as an assimilative context in order to relate what is being learned to what is already known. Third is the idea of activation: the learner must actively relate information in the instruction to existing knowledge--that is, to the assimilative context.

Consistent with the views of Mayer (1980) and other theorists (Weinstein & Underwood, 1985; E. Gagne, 1985), the present study builds upon the view that engaging learners in elaborative processing activities will

Information Processing

A now ubiquitous information processing model (IPM) provides further theoretical basis for conducting research on learning from text. This model has been previously described by a number of authors (cf. Reed, 1982; R. Gagne, Briggs, & Wager, 1988; Mayer, 1984). Derived from the work of Atkinson and Shiffrin (1968) and others, the IPM "... posits internal structures of the human learner and the kinds of processing accomplished by each of these structures (R. Gagne, 1985, p. 70). Figure 2 summarizes the generally agreed upon processes inherent in meaningful information processing.

insert Figure 2 here

Successful encoding appears to be the key to adequately processing and later remembering information (cf. R. Gagne, 1985; Mayer, 1984; Siegler & Klahr, 1982). Craik and Lockhart (1972) stressed the criticality of encoding when they introduced the notion of "levels of processing." According to this view, memory is directly influenced by the nature and degree of encoding operations carried out by the learner (Reed, 1982).

Meaningful, semantic-based encoding is said to produce greater memory than non-semantic encoding (Andre, 1979). According to this view, questions that increase the meaningfulness of information should facilitate learning (W. Wager & S. Wager, 1985). Elaboration--one activity a learner may complete while processing information--is theorized to increase meaningfulness and to support the crucial process of encoding (R. Gagne & Glaser, 1987; Mayer, 1980)

Two critical control processes that are integrated in some information processing models (see R. Gagne et al., 1988). Executive control and expectancies are processes that are said to "...activate and modulate the flow of information during learning" (p. 10). Executive control refers to the learner's self-guided selection and use of cognitive strategies that support encoding and learning. Expectancy refers to the idea that the learner's awareness of the goals and objectives of instruction influence perception, encoding, and subsequent performance.

In sum, the presented framework addresses three superordinate factors involved in the learning process. These are briefly summarized here:

1) **design of stimulus material:** adjunct questions or cognitive strategy activators are design features of the presented instruction that should support learning; other factors such as

facilitate learning. Mayer argues that elaborative activities should produce broad, integrated knowledge structures that are consistent with rule learning. This theoretical perspective predicts that asking questions to promote active integration of the to-be-learned information with existing knowledge will affect encoding so as to improve the learner's ability to apply what was learned in tasks requiring application of intellectual skills.

Expectancy Clarification

A second experimental treatment--the *expectancy clarification* version--includes questions that require subjects to make short answer responses at various points about the goals of the instruction.

Studies by Reder, Charney, and Morgan (1986) and Mayer (1975) reveal that informing the learner of criterion tasks at the start of instruction can positively influence learning. The present study builds upon these findings by considering the relative effectiveness of questioning learners about learning tasks rather than just informing them.

The theoretical basis for questioning students about criterion tasks--while consistent with the attentional aspect of assimilation theory--is primarily derived from forward transfer theory. This view holds that if learners encounter questions of a particular type at the beginning of or within a text passage, they will be prompted to adopt a processing strategy in later passages consistent with the nature of the previously given questions (Sagerman & Mayer, 1987). Learners may continue to apply the adopted reading strategy even when questions cease.

This treatment is also consistent with the work of R. Gagne (1985; R. Gagne, et al., 1988) and others regarding the usefulness of informing learners of objectives in order to establish an expectancy which may serve to guide encoding of the instructional material.

Cognitive Strategy Activation

The third treatment--the *cognitive strategy activation* version--includes directives that suggest that subjects paraphrase key ideas in the instruction. A number of theorists have expressed interest in the use of such directives or "detached cognitive strategy activators" in instruction. Reigeluth (1983), for instance, makes the following comment:

> ...it is not necessary for all learners to have learned a cognitive strategy in order for its corresponding detached activator to improve the quality of the instruction. In fact, since such an activator is so short a piece of the instruction, it may be cost-effective to include it even if just a few students can benefit from it. (p. 213)

Others such as R. Gagne (1986) have indicated that questions designed to induce application of various cognitive strategies is an area that merits exploration. Gagne comments that "...the learner can be asked to *supply* an elaboration by summarizing, by supplying examples, or in some other way" (p. 13).

The planned study considers the relative effectiveness of including adjunct questions in text designed to prompt learners to paraphrase lesson content. Numerous authors have identified paraphrasing as an elaboration technique that should facilitate encoding and learning (cf. R. Gagne et al., 1988; R. Gagne, 1986; Weinstein & Mayer, 1986; Weinstein & Underwood, 1985). Paraphrasing is expected to increase meaningfulness while providing additional cues for later retrieval. A fourth version serving as a control condition does not include any questions.

Subjects

Subjects will include 80 graduate students at San Jose State University in California. Subjects will participate in conjunction with taking an introductory course in educational research methods.

Dependent Measures

The posttest serving as the dependent measure includes questions that are distinct from those presented in the instructional material. These items require subjects to apply selected rules involving the normal distribution and interpretation and calculation of the standard deviation and z scores. Scores will be determined for the total test as well as for the three subcategories of questions: these include 1) calculation of the standard deviation; 2) calculation of z scores; and 3) interpretation of the standard deviation and z scores, respectively.

Qualitative data will also be collected and analyzed after subjects have completed the experimental material. This data will consist of results of eight interviews administered to the highest and lowest performers in terms of total posttest performance in each treatment group and the control group, respectively. A structured interview approach will be utilized to ascertain subjects' perspectives about the various versions of the instructional material, the learning tasks, and the effects of the adjunct questions.

Hypotheses

Hypotheses for the present study concern the relative effectiveness of the three types of adjunct questions in relation to the control condition.

Subjects answering either context activation, expectancy clarification, or cognitive strategy activation questions are expected to outperform subjects in the control condition. It is anticipated that the superior performance will be demonstrated in terms of the total posttest score as well as in terms of performance on the three subcategories calculation and interpretation questions. Figure 3 summarizes the expected effects of the various treatments.

insert Figure 3 here

Directions for Future Research

What considerations should one keep in mind when planning adjunct-question research? This discussion closes with a few possibilities:

1) **learning of intellectual skills:** one can reasonably argue that sufficient research has been conducted involving "recall" as the dependent measure; understanding how questions and cognitive strategy activators may support higher-order learning seems like a more important and useful challenge;

2) **addressing methodological issues:** research should be designed to isolate how questions affect learning; some matters worth documenting in this regard include the amount of time subjects spend completing instruction, the number of adjunct questions that subjects answer correctly, and subjects' ability level in the given subject area. An indicator of ability level could potentially serve as a covariate in data analysis;

3) **employing qualitative approaches:** pure "experimental" approaches to research are no doubt useful; however, interviewing selected subjects after they complete a given treatment may provide additional insights regarding how questions influence the encoding process and help refine subsequent design approaches.

The refinement of questioning techniques in text-based instruction is an intriguing, worthy goal. Through continued research, we may be able to refine the theoretical and instructional design bases required to generate questions that help people learn useful skills in an efficient and meaningful manner.

Figure 1. **A Macro-Level Framework for Adjunct-Question Research**

INSTRUCTION <-->	**LEARNER** <-->	**OUTCOME**
External Events: Stimulus Material	Internal Events: Storage	Memory Structure: Performance
Lesson Structure: - Information - Questions - Feedback - Motivational Design	Information Processing: - Attention - Encoding - Elaboration - Rehearsal	Retrieval: - Procedural/ Declarative Knowledge - Functional Usefulness

---

Figure 2. **Information Processing: An Overview of Elements**

Element	Description
STIMULUS	learning material is designed, developed, released
PERCEPTION	learner encounters and perceives stimulus material through receptors
SELECTIVE PERCEPTION	learner attends to specific elements of the stimulus material
SHORT-TERM MEMORY	information which is the object of selective perception enters this internal structure
ENCODING	learner rehearses, integrates, applies or otherwise processes information in order to store and remember it
LONG-TERM MEMORY	information which is successfully encoded is stored in this internal structure
RETRIEVAL	learner recalls previously encountered information in response to test question or other cue
CONTROL PROCESSES	learner's expectancies about purpose of instruction and use of cognitive strategies while processing information influence encoding

Figure 3. **Hypothesized Influences of Question Types on Learning**

question type	cognitive activity	learning outcome
context activation	extensive elaboration: activation of assimilative context	strong performance on calculation and interpretive questions
expectancy clarification	increased attention and encoding of goal-related content	strong performance on calculation and interpretive questions
cognitive strategy activation	extensive elaboration via paraphrasing: redundant processing	strong performance on calculation and interpretive questions
none	minimal elaboration	poor performance on calculation and interpretive questions

References

Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about questions about what they are reading. In G. H. Bower (Ed.), The psychology of learning and motivation (pp. 89-132). New York: Academic Press.

Andre, T. (1979). Does answering higher-level questions while reading facilitate productive learning? Review of Educational Research, 49(2), 280-318.

Andre, T., Mueller, C., Womack, S., Smid, K., & Tuttle, M. (1980). Adjunct application questions facilitate later application, or do they? Journal of Educational Psychology, 72(4), 533-543.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation, (Vol. 2). New York: Academic Press.

Bing, S. B. (1982). Role of adjunct questions and reading ability levels on rote and conceptual learning from prose. Instructional Science, 11, 129-138.

Carrier, C. A. & Fautsch-Patridge, T. (1981). Levels of questions: A framework for the exploration of processing activities. Contemporary Educational Psychology, 6, 365-382.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior, 11, 671-684.

Gagne, E. D. (1985). The cognitive psychology of school learning. Boston: Little, Brown.

Gagne, E. D., Broughton, S. F., Eggleston, P. D., Holmes, S. E., Hawkins, P. P., & Sheldon, K. L. (1979). Type of post-question and accuracy of concept classification in learning from text. Journal of Experimental Education, 46, 302-306.

Gagne, R. M. (1985). The conditions of learning (4th ed.). New York: Holt, Rinehart and Winston.

Gagne, R. M. (1986). Instructional technology: The research field. Journal of Instructional Development, 8(3), 7-14.

Gagne, R. M. Briggs, L. J., & Wager, W. W. (1988). Principles of instructional design (3rd ed.). New York: Holt, Rinehart and Winston.

Gagne, R. M., & Glaser, R. (1987). Foundations in learning research. In R. M. Gagne (Ed.), Instructional technology: Foundations. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hamilton, R. J. (1986). Role of adjunct questions and subject ability levels on the learning of concepts from prose. American Educational Research Journal, 23(1), 87-94.

Jonassen, D. H. (1985). Generative learning vs. mathemagenic control of text processing. In D. H. Jonassen (Ed.), The technology of text (Vol. 2 pp. 9-45). Englewood Cliffs, NJ: Educational Technology.

Keller, J. M., (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), Instructional-design theories and models: An overview of their current status (pp. 383-434). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mayer, R. E. (1975). Forward transfer of different reading strategies evoked by testlike events in mathematics text. Journal of Educational Psychology, 67(2), 165-169.

Mayer, R. E. (1980). Elaboration techniques that increase the meaningfulness of technical text: An experimental test of the learning strategy hypothesis. Journal of Educational Psychology, 72(6), 770-784.

Mayer, R. E. (1984). Aids to text comprehension. Educational Psychologist, 19(1), 30-42.

Pea, R. D. (1988). Putting knowledge to use. In R. S. Nickerson & P. P. Zodhiates (Eds.), Technology in education: Looking toward 2020 (pp. 169-212). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reder, L. M., Charney, D. H., & Morgan, K. I. (1986). The role of elaborations in learning a skill from an instructional text. Memory & Cognition, 14(1), 64-78.

Reed, S. K. (1982). Cognition. Monterey, CA: Brooks/Cole Publishing.

Reigeluth C. M. (1983). Meaningfulness and instruction: Relating what is being learned to what a student already knows. Instructional Science, 12, 197-218.

Rickards, J. P., & Denner, P. R. (1978). Inserted questions as aids to reading text. Instructional Science, 7, 313-346.

Sagerman, N., & Mayer, R. E. (1987). Forward transfer of different reading strategies evoked by adjunct questions in science text. Journal of Educational Psychology, 79(2), 189-191.

Siegler, R. S., & Klahr, D. (1982). When do children learn? The relationship between existing knowledge and the acquisition of new knowledge. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 2 pp.121-211). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wager, W., & Wager, S. (1985). Presenting questions, processing responses, and providing feedback in CAI. Journal of Instructional Development, 8(4), 2-8.

Watts, G. H., & Anderson, R. C. (1971). Effects of three types of inserted questions on learning from prose. Journal of Educational Psychology, 62(5), 387-394.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd ed., pp. 315-327). New York: Macmillan Publishing Company.

Weinstein, C. E., & Underwood, V. L. (1985). Learning strategies: The how of learning. In J. Segal, S. Chipman, & R. Glaser (Eds.), Relating instruction to basic research (pp. 241-258). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wittrock, M. C. (1984). Teaching learners generative strategies for enhancing reading comprehension. Theory into Practice, 24(2), 123-126.

Title:

An Evaluation of Alternative Distance Tutoring Models for At-Risk Elementary School Children

Authors:

Steven M. Ross
Gary R. Morrison
Lana J. Smith
Evelyn Cleveland

Abstract

Two programs for providing distance tutoring for at-risk elementary school children were implemented and evaluated in successive years. Each program employed 10 college education students as tutors for approximately 25 sixth-grade children. Tutoring exercises were oriented around writing and language skills development. Program 1 was restricted to electronic mail exchanges via a local bulletin board system, whereas Program 2 permitted both electronic mail and teleconferencing via a national network. Results showed moderate to extensive use of both systems by two-thirds of the students. Students' reactions were somewhat negative about Program 1, but positive about Program 2, seemingly because of the greater flexibility of the latter system and stronger relationships established with their tutors. Additional findings and discussion concern the types of communications made, tutor experiences and reactions, advantages and disadvantages of each system, and recommendations for future program design.

An Evaluation of Alternative Distance Tutoring Models for At-Risk Elementary School Children

In an Utopian world, the problem of dealing with differences in learner abilities could be easily resolved by providing each individual with a personal tutor. Aside from the natural appeal of a one-to-one teaching ratio, research on the effects of individualized tutoring has provided fairly strong evidence that it can result in substantial achievement gains by both tutors and tutees (e.g., Bierman & Furman, 1981; Gardner, 1978; Levin, Glass, & Meister, 1987; Slavin, 1988). Unfortunately, in typical educational settings, the realities of limited budgets and large class sizes make such tutoring highly impractical. Even where qualified adults, such as college students, may be willing to tutor children on a voluntary basis (Hirschorn, 1988), significant constraints are the considerable time, expense, and logistical difficulties involved in arranging face-to-face interactions.

These considerations create a rationale for using telecommunications technology to place tutors and tutees in contact with one another from remote locations. Although such "distance" interactions cannot be expected to convey the full impact of face-to-face tutoring, they can seemingly offer many of the advantages with fewer practical constraints (McConnell, 1986). The purpose of the present research was to examine how two distance tutoring models, designed for at-risk elementary school children, operated with regard to degree of use, component activities, and impact on participants.

Although distance learning is associated today with state-of-the-art computer technology, it is hardly a new concept. Correspondence learning, in which tutors and tutees interact through the mail or telephone contacts, have been in existence for many years. Today's computer applications use two basic forms of telecommunications, text teleconferencing and electronic mail (Quinn, Mehan, Levin, & Black, 1983). In teleconferencing, tutors and tutees communicate in "real time" by typing messages on a computer terminal that are transferred by modem to the receiver's computer. The result is a conversation based on immediate exchanges of information, much like that on a telephone. In contrast, electronic mail systems operate in the same way as the exchange of written letters or memos. Once the electronic messages are transmitted to another person, they are stored in a Bulletin Board System (BBS) file (or "mailbox") under his/her name, to be read and responded to at a later time.

For teaching purposes, the advantage of teleconferencing over electronic mail is the ability to conduct an interactive tutoring session that structurally resembles that of face-to-face tutoring. The Cyclops teleconferencing system, used in the Open University in Britain as an instructional supplement in college courses, is one example (McConnell, 1986). Students could "attend" as many Cyclops classes, which were regularly scheduled in place of normal face-to-face sessions, as desired. Evaluations showed attendance and learning under Cyclops to be as high or higher as with conventional tutoring. Electronic mail, on the other hand, has the advantage of creating a relatively permanent written communication that students can access at their convenience

(Quinn et al., 1983). It also provides a means of promoting socialization activities by the sending of written messages to "pen pals" and peers (Sayers & Brown, 1987). The main disadvantages are the delay of communications between teacher and student, and the absence of direct person-to-person contacts (Jones & O'Shea, 1982; Newman, 1987).

The above analysis suggests that teleconferencing and electronic mail both offer unique advantages for distance education. The present research describes and evaluates how each was used with at-risk elementary students participating in the Apple Classroom of Tomorrow (ACOT) program. Alternative distance learning programs employed in successive years are examined. The first (Study 1) involved a locally controlled BBS that restricted communications to electronic mail only, and the other program (Study 2) a nationally-based network system (AppleLink) that permitted both electronic mail and teleconferencing applications.

Study 1

The focus of Study 1, was a distance tutoring model oriented around the exchange of electronic mail messages between tutors and tutees. Activities and attitudes of participants were evaluated through analyses of tutor journals, tutor and tutee survey responses, and the frequency and content of the messages transmitted.

Method

Participants

Tutees consisted of 24 sixth-grade minority students enrolled in the Apple Classroom of Tomorrow (ACOT) class at an inner-city elementary school. There were 15 males and 9 females ranging in age from 10 to 13. The class was one of 13 national sites selected by Apple Computer to establish a computer-intensive environment in which computers would be highly accessible to the children both at school and at home (Baker & Herman, 1988). The present ACOT school was the only site involving distance tutoring and an at-risk student population. Students had been selected to participate in ACOT the previous year on the basis of interest and having access to a home telephone. Because selection did not appear to be systematically biased toward higher ability or higher motivation, the ACOT sample was regarded as reasonably representative of the overall population of sixth-graders at the site school.

Tutors were 10 graduate students (four males and six females) working toward a Master of Arts in Teaching degree at a local university. The tutors' participation was voluntary while offering the incentives of learning about computers and being loaned a home computer system to use during the year. Each tutor was assigned 2-3 tutees during the school year.

Tutoring Methods and Assignments

The host system for the BBS was a 512K Macintosh with a 20-MB hard disk drive and an Apple Personal Modem. One dedicated phone line was used for all incoming calls. Students and tutors accessed the BBS using an Apple IIc computer, an Apple Personal Modem, and Apple Access II software. Once they were on-line, any messages that had been sent to them since their last access were listed by sender's name and message number. They could then select between menu options allowing them to read, send, edit, or delete messages. In sending a message, the user entered the recipient's name and a brief message title. After the message was typed, it could be sent as it appeared, edited or deleted. File transfer was not easily accomplished and therefore was not used.

Considering that electronic mail would be the primary tutoring mode, assignments were oriented around writing skill activities that depended on the exchange of information over time rather than through real-time interactions. The initial assignment asked students to construct a "contrast paragraph" comparing their own background and interests to those of their tutors. Completed paragraphs were posted on the BBS for the tutors to examine and react to through comments and feedback. Later, a printed copy of the assignment was handed in and evaluated by the classroom teacher. Additional BBS assignments during the first semester were patterned on this basic orientation of students submitting early work, receiving feedback, and preparing revised versions to submit in class.

Assignments during the second semester were used to help students expand their reading and writing vocabulary. Tutors were given copies of vocabulary journals that their tutees had created using a data base. Exercises created by tutors listed different groups of words and asked the student to use each correctly in a sentence, identify synonyms, and indicate correct or incorrect uses in completed sentences. Tutors evaluated this work and submitted the grades to the classroom teacher. In addition to these and related formal assignments, tutors and tutees were strongly encouraged to use the BBS informally to share experiences and maintain regular contact.

Instrumentation

Student attitude survey. A 13-item attitude survey was developed to assess student reactions toward three aspects of the BBS: (a) value and utility, (b) helpfulness of tutors and assignments, and (c) technical/operational features. Students reacted to each item using a 5-point Likert-type scale ranging from "strongly disagree" (1) to "strongly agree" (5).

Tutor journals. The tutor journals were forms used by tutors to record any communications received from their tutees. At the top of the form were spaces to indicate the tutee's name, the date the message was sent, and the type of message (social, request for help, etc.). Space below was used to write a description of the message and any impressions about its content.

Tutor survey. At the completion of the program, tutors were administered an attitude survey containing 14 Likert-type items and 10 open-ended questions. The items and questions concerned feelings about their preparedness for the program, the appropriateness of their training and the tutoring activities, strengths and weaknesses of the program, and perceived value of the experience for them and their tutees.

Message data base. A data base was developed for recording and analyzing messages transmitted on the BBS. Categories included number of words, sender, receiver, subject, synopsis of content, and category. For the latter measure, categories were derived on the basis of examining common themes or purposes conveyed in the messages. Eight were identified as inclusive of all messages: Social, BBS Assignment, Tutor Business, Tutee Problems, Miscellaneous, Reminder, Excuse-Explanations, and Grade Report.

Procedure

The tutoring program was initiated in September with orientation meetings for students, teachers, parents, and tutors. These sessions includedbasic demonstrations of how to operate the equipment and use the BBS. Tutors and students met one another at the meeting and also at two social gatherings that were subsequently held at school and the university. The program formally began in early October with the assignment of the first exercise (the contrast paragraph). Tutors gave feedback on the drafts transmitted, and students made revisions until an "acceptable" version was produced. Comparable procedural orientations were used for additional assignments during the year (e.g., word clarification, sentence construction, vocabulary journals).

A hard copy of every electronic mail message recorded on the BBS was obtained for research purposes. Three one-week samples were chosen at random from different periods of the year for analysis. Word counts were made and other critical data were recorded from the messages sampled. The only subjective evaluation concerned categorizing the content of the message. Although some of the messages contained elements of more than one category, it was fairly easy to identify a primary category based on the general intent conveyed (for further details, see Ulrich, 1988).

The student and tutor surveys were administered independently in April. Items were read aloud to students as they read them silently.

Results

How the BBS was Used

Table 1 shows the frequency of messages posted by the different user groups during the three analysis periods. Staff consisted of the

ACOT coordinator and the present authors, all of whom participated in supervising the tutoring project. The results reveal a fairly high level of total responses, ranging from 131 to 231 messages per week. With the exception of boys whose activity rate was very low, the three remaining groups (tutors, staff, and girls) each accounted for close to one-third of the total messages sent. Dividing the frequencies by the number of users yielded mean individual rates of 5.2 weekly messages per tutor (n=10), 14.7 per staff member (n=4), 3.7 per girl student (n=15), and .9 per boy student (n=11). Parallel analyses conducted on the number of messages received showed similar results, with the weekly averages being 4.9 for tutors, 13.1 for staff, 3.3 for girls, and 1.6 for boys. Mean numbers of words per message were 50.7 for tutors, 41.9 for staff, 76.2 for girls and 19.5 for boys.

Insert Table 1 about here

In any given observation week, from half to two-thirds of the girls and boys posted at least one message. Ten of the 26 students, however, failed to post a single message during any of the observation periods. Across the three sample periods, girls (f=167) were almost seven times more active in sending messages than boys (f=24). Girls sent 33 (17%) messages to other girls, but only 2 (1%) to boys; 69 (36%) of their messages were sent to tutors and 87 (45%) to staff. Boys, however, sent half their messages to tutors and half to the ACOT staff, thus sending none to other students. Correlations between standardized achievement test scores and message sending or receiving rates were not significant.

Type of Message

The most frequently used message category was Social (n=274). Topics for these messages included general encouragement for using the BBS, puzzles, chatting about sports, and general friendly letters of correspondence. Messages between students accounted for 12% of the total; the remaining 88% involved tutors and staff communications with the students and with one another.

The second most frequently used category, BBS Assignments (n=104), most often consisted of exercises or feedback posted by the tutors, and responses or questions by the children. The third category, Tutor Business (n=81), entailed messages concerning meetings and other matters such as technical aspects of using the BBS. Tutee Problems (n=40) were messages by tutors concerning students who were failing to complete assignments correctly or on time, and the corresponding replies by staff. As the year progressed, this type of message was used increasingly often, as certain students became identified as chronic offenders who were unlikely to respond to reminders or encouragements. Messages that could not be labeled because of their brevity or content (n=27) were classified Miscellaneous (e.g., "Yes, I will do that"). Other categories consisted of (a) Reminders to students (n=20; e.g., "Remember, your work is due next Tuesday"), (b)

Explanations by students for not responding (n=15), and (c) Grade Reports (n=15) of students' performances.

Student Attitudes.

In response to the attitude survey, students expressed mixed to negative reactions regarding the helpfulness of the tutoring assignments. Specifically, from one-half to two-thirds felt that they were: (a) not understanding corrections made by their tutors (67%), (b) not learning much from their tutors (62%), (c) receiving little help with their writing skills (65%), (d) receiving insufficient time from their tutors (56%), (e) finding the BBS assignments difficult (57%). Only a little over half felt that they had enough on-line time (the 30 min. allotted per access) to complete messages, and about one-third found that the line was usually busy when they tried to call.

Correlations were computed between the number of messages received/sent and survey responses. As might be expected, those who were more active in sending messages were more likely to prefer writing social messages than tutoring assignments (r=40, p< .04), and were more bothered by (a) the public display of the messages written (r=.39, p < .05) and (b) the difficulty of accessing the BBS (r= .42, p = < .04). More active users were less inclined to feel that their tutors were helping them with their writing skills (r = -.45; p < .03). The relationships seem largely due to the very active users being more involved in social than in academic communications. Girls had a higher preference than boys for social messages and were less satisfied with the BBS access and the helpfulness of their tutors (p's < .05).

Tutor Attitudes

Tutors' responses (n=9) to the Likert-type survey items indicated that most felt sufficiently trained in computer skills (78% agreement) and in the subject matter (78%) to do the work expected. There was less agreement, however, regarding the sufficiency of the orientation training in acquainting them with teachers' objectives (22% agreement), teachers' expectations (44%), and students' expectations (11%). Nearly all tutors (89%) felt that a more intensive training program was needed; two-thirds (67%) would have liked additional opportunities for personal contacts with students.

The majority of tutors (56%) felt that they related well to their tutees, whereas close to half (44%) felt that their tutees related well to them. Two-thirds (67%) agreed that they were given sufficient materials to plan the assignments. Less than one-fourth agreed that teachers communicated with them adequately (22%) or provided sufficient feedback on their roles as tutors (22%).

When asked what types of BBS assignments were appropriate for the type of delivery method, nearly all the tutors indicated a preference for the first assignment type (friendly letter writing). One tutor responded that the writing skill development exercises "took advantage of the computer's capabilities while the fill-in-the-blank or multiple-choice vocabulary exercises should have been on paper and given to the students." However, another volunteered that writing skills were too difficult to communicate during the 30 minutes of allotted BBS access time. The general consensus was that with a more flexible BBS system most of the assignments would have worked well.

The greatest strengths of the program were perceived as providing participants with exposure to new technology and software; creating an awareness of the special learning and communication needs of the children; providing support to students; and, in some cases, allowing the development of a friendship via the BBS. One of the most important strengths was perceived to be writing skills development, as reflected in the following tutor comment: "The strength of BBS had to do with the fact that the children had to rely on written communication to get their work done. They were forced to read and comprehend as well as write in a way that others could understand." Weaknesses were identified as (a) lack of effective communications with teacher and tutee; (b) delayed feedback; (c) insufficient phone lines; (d) inability to talk on-line; (e) inability to print out assignments; (f) tutees not completing the assignment on time; (g) inappropriate assignments that could have been better handled by printed material; and (h)a lack of planning time.

When asked what was expected of them, the tutors listed motivating students, helping with assignments, correcting assignments, and interacting socially. Several tutors, however, indicated some confusion about what their primary roles should be. All but one tutor was highly positive about recommending the ACOT tutoring experience to future prospective tutors. The one negative response was that the responsibilities demanded too much time. All tutors felt that they could have tutored better in person at times because the BBS was very impersonal and assignment corrections often became to wordy. Other comments expressed concern over the apparent lack of parental support. One tutor wrote, for example, "No wonder Jill does not respond to my messages or worry about F's on BBS assignments, her parents obviously don't feel that the BBS is of enough importance to let her hook it up to the phone line." Whether due to parental behaviors or other factors, the failure of some students to complete the assignments or communicate regularly was the biggest source of frustration for the tutors as a group.

Based on the results, positive aspects of the program were viewed as the high level of BBS usage by most of the students, and the perceived potential to channel that activity into academically-focused exercises. A significant impediment to program activities was the operational limitation of the single-line BBS system which made access highly difficult, restricted the time and space available for messages, and did not allow file transfer or real-time communications. Many students also found the assignments difficult and not helpful for learning. In view of these outcomes, goals for the second year were to increase the effectiveness of the program through the use of a more versatile BBS and more carefully planned assignments.

Study 2

The second-year project and evaluation were comparable to those of Study 1 with the following major differences: (a) AppleLink, a national network system that permited both electronic mail and teleconferencing, was used as the BBS; and (b) due to a delays in establishing BBS accounts and allocating computer equipment, the project was not initiated until January of the school year.

Method

Participants

Students were 27 sixth-grade minority students (9 boys and 18 girls) from the same school as in Study 1. Tutors were again 10 education students from a local university (8 females and 2 males). Based on the Study 1 results, an attempt was made to establish closer student-tutor relationships by having each tutor spend three days at the school, in October, to become acquainted with his/her tutees and the school environment.

Procedure

Due to the late start-up time, fewer exercises could be assigned relative to Study 1. One assignment was a "friendly letter" that the students composed and sent to their tutors who gave suggestions for improvement. The second exercise was the construction of a personal time-line of each student's life. A third was the writing of an autobiography using the time-line as a reference. A fourth was a contest in which students earned team points for completing assignments and for having on-line conversations with special people (teachers or staff) who accessed the system intermittently. As in Study 1, students were also encouraged to use the system frequently and regularly to communicate with others. Such conversations took the form of both on-line conversations and electronic mail exchanges.

Data regarding program activities and reactions were obtained from instruments adapted from Study 1 and several additional measures. These included: (a) student and tutor attitude ratings and open-ended questions, (b) student interviews, (c) tutor journals, and (d) AppleLink billing records indicating individuals' frequency of access and accumulated on-line time.

Results

How the BBS was Used

Due to the inability to obtain copies of actual messages under the AppleLink system, data regarding uses of the BBS were obtained from two secondary sources, tutor journals and billing records. The journals provided a record of the messages sent by students to their tutors during March 1 to April 15, a six-week period when all participants had active

BBS accounts. Tabulation of the messages recorded showed that students sent tutors a total of 97 messages. This represented a 41% increase over the 69 student-tutor messages sent during the six-week period monitored during Study 1, but comparability of the data is weakened by differences in school and tutoring activities occurring at those times. The messages were categorized as being one of five types: Personal/Friendly (n=39), Tutoring Assignments (n=26), Mail Responses (n=18), School Activities (n=11), BBS Operations (n=3). The length of the messages ranged from 10 to 90 words, with an average of about 25 words.

The billing records indicated the total amount of time each user was on line each month from January through April. A summary is provided in Table 2. As revealed, virtually all tutors and most (approximately two-thirds) of the tutees accessed the system each month. Tutor on-line times greatly fluctuated across time periods, a direct result of the planning and communication requirements for different tutoring exercises. Student on-line times were more consistent averaging close to one hour per week overall. The distribution of student on-line time, however, was highly skewed, with some individuals spending considerable time (e.g., as high as 71 hours in one month) and others only a few minutes.

Student Attitudes

Survey results. The attitude survey was completed by the 23 students who attended class on the day of its administration. In direct contrast to Study 1, responses were the most positive on items concerning relationships with tutors: (a) liking one's tutor (100% agreement), (b) wanting to spend more time with tutors (87% agreement), and (c) feeling that tutors liked them (70% agreement). Nearly three-fourths (74%), however, felt that they did not have enough contact with their tutors. Students also compared the tutor interactions favorably with those involving peers, with the majority feeling that they: (a) preferred tutors messages to friends messages (56%); (b) preferred talking to tutors than to friends (63%); (c) would rather have their present tutor than one from their school (61%), or (d) preferred leaving messages for tutors than for friends (65%). With regard to assignments, 69% agreed (with the remainder undecided) that they liked assignments on the BBS, and 65% agreed that they were learning from their tutors. Their main frustration was not having the opportunity to use the BBS for as much time as they wanted (82% agreement).

On open-ended items, tutees were asked to identify what they liked most and least about the tutoring experience. Sixty-five percent felt that it gave them a chance to meet new people and learn new things. One tutee wrote, "Getting to know someone from a college and that my tutor is helping me to improve my work." Another responded, "I liked it because it gave me a chance to learn something new." Students also liked being able to talk to friends and "getting to know people all over the world." The consensus regarding the least favored aspect was the limited on-line communication with tutors.

Interview responses. Interviews were conducted over a three-day period the first week of May. All 24 students who attended school during that period agreed to the interview. On the average, students reported spending 5.4 hours a day on the computers at school compared with an average home usage of 1.2 hours. When asked what they liked best about their ACOT experiences, they cited meeting new friends ($n=8$), learning new activities ($n=3$), having new and much more software ($n=2$), etc. Six (25%) students identified the use of AppleLink as their favorite experience. When asked what was liked least about ACOT, the only four students who responded each mentioned "more work" in their replies.

Approximately one-third had not used AppleLink from their home but did so from school. Reasons were equipment problems, no access to a phone, or not knowing how to operate the system. Each student was also asked to describe his or her use of AppleLink with regard to specific applications of the Post Office (send/receive mail), ACOT Chat Room (ACOT-only teleconferencing), or the Apple Lobby (general public teleconferencing). The Apple Lobby was used by more than 92% of the students. The main reason for its use was the adventure of meeting new people from across the country. Less than 25% of the students, however, used the ACOT Chat Room more than once or twice; those who did not use it responded that "there was never anybody there." Fifteen of the 24 students indicated they used the Post Office for sending/receiving messages. When asked how the tutoring experience could be improved, 96% felt that more personal contact was needed. Specific suggestions were to have occasional meetings of the participants at school or to arrange set times to "meet" in the ACOT Chat Room.

Tutor Attitudes

All 10 tutors either agreed strongly or agreed that they were competent to tutor based on their academic skills and computer expertise. Seven indicated that the assignments were appropriate for the BBS, and all agreed that AppleLink was easy to use. Six felt that they related well to their tutees.

At least half of the tutors disagreed, however, that they received adequate communication from classroom teachers (80%) or that they were able to communicate their expectations to their tutees (50%). The majority were undecided on whether they had enjoyed the tutoring-mentoring process (60%) and on whether the tutoring assignments were consonant with the classroom teacher's assignments (80%). None indicated that they were disappointed with their tutees' responses.

When asked to list positive aspects of the program, many tutors identified the initial face-to-face meeting, students' introduction to new technology, the motivational benefits of using a computer for learning, and the computer skills they personally gained. The most often mentioned negative aspect was the lack of coordination between the tutors, tutees, and school teachers. One tutor responded, "I felt isolated. There seemed to be no strong central advocate for the AppleLink activities in the classroom who had enough time to see that things got done." Other negative aspects included the lack of personal contact with students, or failure to communicate or complete assignments by some students, and equipment failures.

Five of the tutors (50%) felt the tutoring would have been more effective in person; mainly, because the students were so "elusive." Four of the tutors stated that a combination of distance tutoring and face-to-face personal tutoring would be best. Eight tutors indicated that their tutees were frequently unresponsive. One wrote, for example, "I tried to be funny and supportive and show my interest in their lives, but the replies I got were very short and didn't develop into any kind of dialog." Generally, the tone of the open-ended comments seemed largely dependent on the responsiveness of their tutees, which in at least one-third of the cases was extremely low. Thus, some tutors described their tutees' activities and productivity in positive terms while others were negative. Interview responses reinforced these impressions, while also providing suggestions for improving the program. These included having workshops for using the BBS; holding students more accountable for their work; and having more teacher support; more structure, greater integration between classwork and BBS assignments, and more personal interaction. Reactions to the prospect of participating in similar programs were positive, but qualified on the basis of the above types of changes being made.

Discussion

Evaluation of the two distance tutoring programs revealed common outcomes as well as some contrasts. In both instances the potential benefits of the tutoring activities for improving learning and motivation were clearly recognized and acknowledged by staff and tutors. Involving students in reading and writing activities as a natural means of communicating on the BBS was seen as an especially valuable feature. Another consistent finding was the general enjoyment and extensive use of the BBS by most students. Such was particularly the case with the AppleLink system (Study 2) which permitted students to converse in real time in addition to exchanging electronic mail messages. In both programs, about one-third of the students were highly active BBS users, accessing the system nearly every day, while another one-third were almost completely inactive. The main reasons for the latter (especially in Study 2) appeared to be equipment problems, and lack of motivation to complete assignments or to use the system in general.

Differences between program outcomes generally reflected the contrasting environments that were created by each. In the case of the local BBS (Study 1), users were greatly frustrated by the difficulty in accessing the system as well as time delays and the restrictiveness of electronic mail communication (c.f. Newman, 1987). In the case of the AppleLink system, access was almost always immediate, but students appeared to be diverted by many of the extraneous system features (e.g., public conversation rooms, and information sources), resulting in less commitment to the academic activities assigned. While the greater flexibility of the AppleLink system expanded ways of communicating, it also created increased training demands. For example, both students and tutors experienced difficulty in learning how to upload and download files, as was required for several of the exercises.

Based on the findings, a number of guidelines are suggested to improve the effectiveness of future programs. First, greater front-end planning is required to more closely match program activities to curriculum objectives and media capabilities. Although most of the tutoring exercises assigned generally appeared appropriate, their designs were often dictated primarily by the immediate demands of maintaining a consistent program of activities as opposed to being derived from a systematic analysis of what instructional needs each would support during the school year. There were also some activities (e.g., vocabulary exercises) that would have been more appropriately conveyed through conventional delivery modes (printed exercises or classroom instruction) than by distance tutoring. Compared to the exercises based on electronic mail exchanges, those that took advantage of the on-line interactive capabilities of the AppleLink system were not as well developed or frequently used.

Second, tutoring exercises need to be documented for participants in advance of their assignment, with objectives, completion dates, activities, and products clearly specified. In the present programs, tutors often received only brief oral descriptions of exercises a few days in advance of the start-up time. This left them little time to prepare materials and uncertain about expectancies.

Third, in order to take full advantage of distance tutoring capabilities, students and tutors need to be fluent with the operational procedures for the system. In the case of the present AppleLink program, insufficient training on file operations (uploading/downloading, printing, etc.) resulted in lack of preparedness by students for completing several of the assignments independently.

Fourth, and perhaps most importantly, increased personal contacts between tutors and tutees appear crucial for establishing meaningful relationships and strengthening the social aspects of the learning experience (Jones & O'Shea, 1982; Paulet, 1987). Accordingly, the need for more frequent face-to-face meetings between students and tutors was one of the most consistent and strongest recommendations made by participants. Some support for this idea is provided by the Study 2 results which suggested the establishment of closer tutor-tutee relationships and more positive student attitudes relative to Study 1. A strong contributing factor appears to have been the three full-day classroom visitations that Study 2 tutors made at the beginning of the year.

Overall, the results of the two studies reinforce the notion that distance tutoring is attractive to students and can involve them in useful learning activities using a combination or separate applications of electronic mail and teleconferencing. However, as in the case of conventional tutoring, success does not directly follow from putting students who need help in contact with capable tutors (see review by Cohen, 1986). Tutoring activities must be carefully planned with regard to individual needs and educational objectives, while positive interpersonal relationships of mutual trust and caring need to be developed. Unlike conventional tutoring, an additional requirement for distance tutoring is to establish these conditions within the constraints and unique environment imposed by the technological systems. This design integration process appears likely to offer instructional designers the greatest degree of challenge.

Table 1. Frequencies and Proportions of Messages Posted by Different User Groups in Study 1

Period	Tutors (n = 10)		Staff (n = 4)		Girls (n = 15)		Boys (n = 11)		Total (n = 40)	
	f	p	f	p	f	p	f	p	f	p
1	72	.31	93	.40	53	23	13	.06	231	1.00
2	42	.32	44	.34	37	.28	8	.06	131	1.00
3	42	.20	83	40	75	.36	5	.03	205	1.00

Table 2. Percent Participation and Mean One-Line Time for Tutors and Tutees in Study 2

Period	Tutors (n = 10)		Students (n = 27)	
	% Using	M Time (hrs.)	% Using	M Time (hrs.)
Jan.	80	17.8	59	4.4
Feb.	100	3.0	67	4.3
Mar.	100	10.1	62	2.8
Apr.	100	3.9	78	5.3

References

Baker, E., & Herman, J. (1989). The ACOT report card: Effects of complex performance and attitude. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Bierman, K. L., & Furman, W. (1981). Effects of role and assignment rationale on attitudes formed during peer tutoring. Journal of Educational Psychology, 73, 33-40.

Cohen, J. (1986). Theoretical considerations of peer tutoring. Psychology in the Schools, 23, 175-186.

Gardner, E. W. (1978). Peer assistance through tutoring and group guidance activities. The Urban Review, 10, 45-54.

Hirschorn, M. W. (1988, January). Coalition of 120 colleges hopes to encourage a million students to tutor "at-risk" youths. Chronicle of Higher Education, pp. 35, 38.

Jones, A., & O'Shea, T. (1982). Barriers to the use of computer assisted learning. British Journal of Educational Technology, 13, 207-217.

Levin, H. M., Glass, G. W., & Meister, G.R. (1987). Cost effectiveness of computer-assisted instruction. Educational Review, 11, 50-72.

McConnell, D. (1986). The impact of the Cyclops shared-screen teleconferencing in distance tutoring. British Journal of Educational Technology, 17, 41-74.

Newman, D. (1987). Local and long distance computer networking for science classrooms. Educational Technology, 27, (6), p. 20.

Paulet, R. (1987, September). Counseling distance learners. Tech Trends, 32, 26-28.

Quinn, C., Mehan, H., Levin, J.A., & Black, S.D. (1983). Education in non-real time: The use of electronic message systems for instruction. Instructional Science, 22, 313-327.

Sayers, D., & Brown, K. (1987). Bilingual education and telecommunications: A perfect fit. The Computing Teacher, 14, 23-24.

Slavin, R. E. (1988). Educational psychology: Theory into practice (2nd ed.) Englewood Cliffs, NJ: Prentice-Hall, Inc.

Ulrich, L.M. (1988). A study of electronic bulletin board usage by participants in the Memphis ACOT Program. Unpublished Master's Thesis, Memphis State University.

Author Note

This paper was originally published in *Computers and Human Behavior*, *6*, 247-259.

Title:

The Development of Writing Skills in a Computer-Intensive Environment

Authors:

Steven M. Ross
Lana J. Smith
Ellen Woodson

Abstract

Uses of the computer to support the development of writing skills was examined in a computer-intensive educational context, the Apple Classroom of Tomorrow (ACOT). Subjects were 55 sixth-grade minority students, 25 of whom were ACOT participants and the remainder members of a matched control group attending the same school. Computer-based writing experiences consisted of using word-processing as a primary classroom writing mode and electronic mail as a supplementary activity. Results from a year-long study indicated high enthusiasm and preference by ACOT students for word-processing applications, which they used naturally and routinely for virtually all writing tasks. The electronic mail component was regarded as less effective due to hardware problems and the lack of an organized activity plan. Evidence for the achievement benefits of the computer-based writing program was provided by the superiority of the ACOT group over the control group on writing samples collected during the school year.

The Development of Writing Skills in a Computer-Intensive Environment

Using computers to teach and develop writing skills a topic of considerable research interest. In the specific case of word-processing, the important advantages are viewed as facilitating the revision and editing of drafts, and improving student attitudes toward the writing process (Collins, Bruce, & Rubin, 1982; Daiute, 1985a; 1985b; Heebner, 1988; Quellmalz, 1989). MacArthur (1988) and others (Daiute, 1985a; Calkins, 1986) have further emphasized the potential impact of the computer on the social context for classroom writing.

Although many authors and researchers agree that use of a word-processor can improve attitudes and confidence about writing (e.g., MacArthur, 1988; Madian, 1986; Newman, 1984; Rodrigues, 1985; Strictland, 1986; Sudol, 1985), there is little evidence from well-designed studies that indicate its impact on writing improvement. Rather than trying to establish the computer's "effect" on writing improvement (see Clark, 1983), it would seem more productive to investigate ways in which computer-based writing activities can be effectively integrated with the curriculum and conventional classroom activities. One type of activity consists of "computer-specific" applications, such as electronic mail or teleconferencing (Heebner, 1988; Ross, Smith, Morrison, & Erickson, 1989), which extend the forms of writing that can be achieved with conventional (paper and pencil) methods. Another entails "computer supported" applications (such as word-processing) in which the normal writing curriculum and teaching methods can be maintained, but the computer is used as a primary classroom writing medium. Unfortunately, insufficient computer resources at most schools severely limit students' access to computers and thus the opportunities to implement substantive writing programs of either type.

In the present research, the establishment of a unique elementary school environment, the Apple Classroom of Tomorrow (ACOT), permitted examination of long-term uses of both computer-specific and computer-supported writing applications. ACOT comprises an experimental computer-saturated environment in which each student receives a personal computer to use during the school day and another one to use at home. We were thus able to examine students' writing activities under conditions that anticipate future educational environments in which computers will be highly accessible as a support tool both in and out of school. In addition to using word-processing as a primary writing mode, students communicated with one another and with college student tutors via written messages transmitted over an electronic Bulletin Board System (BBS).

A distinctive aspect of the present ACOT classroom relative to 15 others nationwide was its location in an inner-city school attended almost exclusively by minority, "at-risk" students. The availability of a matched control group attending the same school permitted a quasi-experimental design to be employed in assessing program influences on writing performance. However, in view of the complexity of the various ACOT interventions (writing related and otherwise), and the associated threats to internal validity, the group comparisons were supplemented by extensive qualitative and quantitative data from interviews, surveys, and observations of ACOT participants involved in the writing component. Major research questions addressed were as follows:

1. How is writing taught and integrated with other learning activities in the computer-saturated, ACOT environment?

2. What are students' attitudes toward practicing writing using word-processing and electronic mail?

3. How do teachers view the strengths and weaknesses of the ACOT experiences for developing writing skills?

4. Do the ACOT writing components appear beneficial for improving writing quality relative to conventional methods?

Method

Subjects

Subjects were 55 sixth-grade minority (Black) students attending the same inner-city elementary school in Memphis, Tennessee. Twenty-five students were members of the sixth-grade ACOT class and 30 were members of a conventional (control) class. Based on their backgrounds, home environments, and academic records, nearly all of these students can be considered academically at risk (see Slavin & Madden, 1989). Analyses of scores on the reading vocabulary, reading comprehension, and language mechanics subscales of the California Achievement Test (CAT), administered in the fall and spring of the fifth-grade school year, failed to establish either the ACOT or the non-ACOT classes as superior (Kitabchi, 1988). Based on these results and informal teacher evaluations, neither group appeared to have an advantage over the other in writing ability at the start of the study. However, because assignment to classes was not completely at random, differential selection (Campbell & Stanley, 1963) cannot be ruled out as an internal validity threat to the group comparisons.

Design

This study was a descriptive and quasi-experimental evaluation of the writing activities and skills achieved in a computer-saturated elementary school class. Outcomes examined included (a) the nature of writing activities in the ACOT classroom; (b) experiences and attitudes of teachers, students, and tutors regarding those writing activities; and (c) differences between ACOT and control students in writing skills. These outcomes were assessed by a variety of quantitative and qualitative measures including interviews, written surveys, classroom observations, and writing samples collected from students during the year. Although the experimental and control groups appeared to be comparable in student abilities and characteristics (see above section), and followed the same basic sixth-grade curriculum, they were taught in different classrooms and by different teachers. Thus, many extraneous factors, aside from type of writing program used (computer-based or conventional), could have contributed to any differences in performance. The writing skills analysis was therefore regarded primarily as a source of suggestive evidence about program effects, not as a basis for supporting causal interpretations.

The ACOT Classroom

ACOT was initiated in 1983 as an experimental partnership between Apple Computer, Inc. and 13 public school classrooms in five states. The rationale was to create "innovative learning environments" that take full advantage of technology, study how these environments affect learning and teaching, and promote positive changes based on the results (Baker & Herman, 1989; Kitabchi, 1987, 1988). The common feature across sites was the availability of computers and software for each student to use in the classroom and at home. Each of the sites, however, was given freedom to develop its own program goals and strategies for achieving them.

For the present ACOT site, the overall project objective was to improve the basic learning skills of its at-risk student population. One of its program activities was the establishment of the electronic mail component to link students with tutors at a local university (Ross, Smith, Morrison, & Erickson, 1989). The tutors assisted students with homework assignments, assigned supplementary work, and provided personal encouragement. The electronic bulletin board system (BBS) used for the mail exchanges was operated locally by an ACOT staff member. A detailed description and evaluation of the BBS component is provided in Ross, Morrison, Smith, & Cleveland (1989).

Additional program emphases were supplementing conventional instruction by encouraging teachers to use telecommunications, multimedia, and curriculum software tools to enrich learning of basic subjects. As will be described below, the most pervasive of these activities was regular classroom and home use of word-processing (using the AppleWorks word-processor) for school writing assignments.

Classroom Writing Activities

What mainly distinguished the ACOT writing activities from those of conventional class was the students' extensive, routine use of word-processing to complete homework and in-class writing assignments. Homework in various subjects was completed on the word-processor at home and printed out at school. Specific writing assignments included journal reports, letters to pen pals, and English compositions. On occasion, exercises oriented around writing skills covered in the textbook were devised for practice on the word processor. Students had composition for one 40-minute period a week and were allotted 20 minutes a day for specific writing assignments. During the day, however, students usually had access to their computers to complete regular classroom assignments in different subjects. The Tennessee Instructional Mod . (TIM), a direct instructional teaching orientation, was followed for instruction in writing and other subjects.

The electronic mail program was also intended for use in developing language skills and vocabulary through the completion of exercises assigned by the students' tutors. These assignments, however, were sporadic and few in number. Greater emphasis was given to involving students in using the BBS writing to communicate informally with their friends and tutors. For these exchanges, the content of the messages was stressed; no corrections of spelling or grammatical errors were expected or typically provided.

Although the control students had fairly extensive contact with computers as a result of taking a one-hour computer literacy class each day for the entire year, they did not work with computers during regular class time. Rather, they were taught in the traditional manner, using the same language arts curriculum and objectives as the ACOT students, but with pencil-and-paper materials exclusively. As in the ACOT class, the TIM orientation was followed in presenting lessons. In general, ACOT and control classes appeared to be fairly comparable in (a) allocated time for writing, (b) basic teaching methods (e.g., TIM), (c) learning materials other than computer-based resources, and (d) curriculum objectives. Major variations consisted of the assignment of groups to different classrooms and teachers, and the ACOT group's use of a word-processor for completing most of their written assignments and electronic mail to communicate with peers and tutors.

Instrumentation and Data Collection

Writing assessments. Writing assessments were administered in January and April of the school year to evaluate the writing skills of the ACOT and control classes. Both samples were written by both groups using paper and pencil. In the first (January) writing sample, the students wrote a paragraph describing an apple. The exercise began with the teacher asking students to "brainstorm" about its possible uses, importance, features, and so on. The students were then instructed to write their paragraphs, paying attention to organization, grammar, and spelling. The second writing sample (April) was a letter to the school principal addressing the issue of whether students should have individual computers at their desks.

The paragraphs were evaluated without knowing the writer's identity or treatment group by two English teachers. The criteria were based on skills commonly taught throughout the communication acts curriculum adopted by the state and local education system. The first four entries consisted of: (a) counting the number of words, (b) counting the number of sentences, (c) identifying if there was a topic sentence (yes = 1 or no = 0), and (d) counting the number of details in the paragraph. The topic sentence measure specifically involved assessing whether the sentence in question was supported by information or details in subsequent subordinate sentences in the paragraph. Details, in turn, were defined as points, facts, phrases, or explanations defining or elaborating the topic conveyed. Details were counted in each sentence independent of the presence or absence of a topic sentence or of their relationship to the topic sentence.

The second group of criteria addressed grammar. The specific variables are listed below with a brief explanation of how they were assessed.

1. Number of misspelled words. Each misspelling was counted as one error, regardless of whether the identical error previously occurred. A word correctly spelled but having the wrong tense or meaning was counted as a mechanical error (see "6" below) not as a spelling error.

2. Number of sentence fragments. Each fragment that appeared as part of a complex sentence or as a separate simple sentence was scored as one error.

3. Number of run-on sentences. Each grouping of contiguous sentences strung together as one run-on sentence was scored as one error.

4. Number of capitalization errors. Capitalization errors included errors of commission as well as omission. Handwriting, however, was frequently a problem in identifying whether upper- or lower-case letters were used. When such questions arose, the two raters examined the character in question and reached a mutual decision.

5. Use of pronouns. Use of pronouns was evaluated subjectively and globally as "inappropriate" (score = 0) or "appropriate" (1) for producing a clear and effective paragraph.

6. Mechanical errors. Mechanical error scores involved tabulating errors representing: (a) the misuse or omission of articles, (b) incorrect use of the singular or plural form of a word, (c) subject/verb agreement, (d) tense agreement throughout the paragraph, and (e) incorrect punctuation. In addition, redundancy of words or ideas was evaluated by a dichotomous satisfactory (1) or unsatisfactory (0) rating.

Using a five-point scale ranging from "poor" (1) to "excellent" (5), each paper was evaluated holistically for its unity and coherence in expression. The same evaluation criteria were used with both writing samples, but the letter to the principal (Sample 2) included an additional score indicating whether or not the students used a correct letter form (Yes or No).

The two evaluators discussed each variable to ensure that they had the same understanding of the operational meaning and scoring procedure. After individually scoring several of the same writing samples, the evaluators compared their scores and clarified their definitions of terms and methods. Interrater reliabilities, assessed by correlating the raters' scores on each variable, were quite high, with median r's of .93 and .98 on Samples 1 and 2, respectively.

Student attitude survey and interview. A 14-item attitude survey was administered to ACOT students to assess their reactions toward their writing experiences. The items consisted of statements to which levels of agreement or disagreement were indicated on a five-point Likert-type scale (5= "strongly agree"; 1= "strongly disagree"). Topics covered included keyboarding, word-processing, electronic mail, and interactions with tutors. Students also participated in a 5- to 10-minute interview regarding their experiences with ACOT and the writing component.

Tutor survey and interview. A survey consisting of seven open-ended questions was administered to the 10 tutors to assess their perspectives of ACOT writing activities, specifically in reference to the electronic mail component. Questions concerned (a) students' writing skills and achievement, (b) the strengths and weaknesses of the program for teaching and promoting writing, (c) tutors' contributions to improving writing, (d) tutors' preparation for their assigned activities, (e) the appropriateness of assignments, and (f) suggestions for improvement.

Data Collection Procedure

In January of the school year, three instruments were administered. First, the student attitude survey was administered to students in the classroom. Items were read aloud as students read them silently. Second, a structured interview was conducted with the ACOT teacher and the ACOT coordinator. Third, the first writing sample was obtained from the ACOT and control students. Students wrote the composition in class and submitted it to the experimenter when done. In April, the second writing sample was written by students during a regular class period. Students were also interviewed in person during free periods.

Results

Teacher and Coordinator Interviews

The ACOT classroom teacher was interviewed regarding the nature and perceived effectiveness of the writing experiences. Both were extremely positive about the word-processing and electronic mail components but at the same time, frustrated by all that needed to be done and the lack of an established activity plan to follow. The teacher described the students' primary uses of the computer as involving the word-processor. Specific writing assignments included essay questions and English compositions assigned for homework. The students did not have any specified free-time to work on the word processor but were allowed "adventure time" after their other assignments were completed to either work on the word processor or on other software. The teacher strongly believed that the word-processor simplified the work of making corrections and improvements which, in turn, improved students' writing skills and interest. The word-processor, overall, was characterized as an "invaluable" tool.

Among the perceived limitations of the program was its lack of carefully planned writing activities to develop writing skills. Another was that not all students appeared to find electronic mail communications appealing and that only the more expressive students used it extensively. There appeared to be some problems in integrating electronic mail effectively with other work due to its newness and experimental nature; also a problem in isolated cases was a "communication gap" between the college student tutors and the sixth-graders. The teacher also expressed interest in having a greater number of CAI writing programs to accommodate individual differences. Although there was much to do in coordinating writing activities with all of the other ACOT and regular teaching responsibilities, little extra work or preparation was needed to implement word-processing (AppleWorks) as a primary writing mode. Every student learned to use the word-processing program for composing and printing drafts within several days of intermittent practice and employed it for nearly all in-class and homework assignments thereafter. Over time, they developed greater skills with special word-processing applications such as centering text, deleting sections of text, numbering pages.

Student Survey

Student survey items assessed reactions to electronic mail communications and to word-processing. Responses generally indicated dissatisfaction with the electronic mail assignments and their helpfulness to writing improvement. Specifically, the majority of students indicated that they: (a) did not understand their tutor's corrections of their work (67% agreement); (b) were receiving little help with their writing skills (65%); and (c) would prefer to use the BBS to write messages to friends rather than to their tutors (76%). Students were almost evenly divided on the questions of whether they preferred being tutored via electronic mail messages over face-to-face tutoring, and whether they preferred reading messages over writing messages (see Ross, Morrison, Smith, & Cleveland, 1989 for a more detailed report).

Reactions to the word-processor were much more positive. All students (100%) indicated that they liked typing their work, and nearly all (92%) preferred doing assignments on the computer rather than by hand. Aspects of word-processing that were identified by students as especially desirable were the editing features, seeing one's work printed out, and typing as opposed to writing by hand.

Tutor Interviews on the Electronic Mail Activities

Strengths and weaknesses of ACOT for developing writing skills.

Specific strengths of the electronic mail activities as perceived by the tutors were practice and repetition (n = 7), the requirement to communicate through writing

(n = 2), reinforcement of vocabulary and spelling (n = 2), facilitation of editing and improvements (n = 2), and the establishment of favorable contexts for writing (n = 3). Weaknesses were identified as the inability to give adequate feedback on the current BBS system (n = 5) and the lack of structure to the writing activities (n = 4). Other responses noted the difficulty of accessing (n = 3) and communicating over the BBS (n =3), the limited program duration (n = 1) and limited number of assignments (n = 1).

Tutor's personal contribution. Nearly all tutors (n = 7) felt that they failed to make a significant contribution to their tutees' writing skills (for the reasons indicated on the previous question). Several felt that they were somewhat helpful due to providing positive role models and establishing social relationship with their tutees. Despite accomplishing less than was hoped, most of the tutors (78%) felt that they had sufficient background for tutoring. A frequently expressed concern (56%) was the lack of orientation received regarding the students' needs and level of achievement.

Appropriateness of assignments. Approximately half of the tutors felt that all assignments were appropriate, whereas the other half felt that only some of the assignments (e.g., vocabulary exercises) were appropriate. Exercises identified as the most beneficial for improving writing skills were: (a) students writing paragraphs or letters and sending them to tutors, (b) students writing sentences using given vocabulary words or their own selections, (c) students creating stories by sequencing given sentences, (d) students "playing teacher," (e) tutors writing sentences for students from words that students provided, and (f) students rewriting sentences to correct grammar or change meaning. The majority of tutors (n = 6) indicated that the main need was for increased opportunities to practice writing, both in regular classroom assignments and over the BBS.

Writing Samples

Table 1 presents the group means and t-test results for each of the writing skill evaluated in Sample 1. As revealed, significant differences ($p < .05$) were obtained on four variables. The ACOT students wrote more, used more topic sentences, and had fewer mistakes in using singular and plural verbs. The control students, however, made more appropriate uses of pronouns.

Insert Table 1 about here

Table 2 presents results for writing Sample 2. The ACOT group had significantly ($p < .05$) fewer errors in spelling, capitalization, punctuation, and subject-verb agreement, and had better overall content. For the latter evaluation, the general impression was that the ACOT letters were more focused in addressing the topic and featured less extraneous information. Such is reflected in the directionally (but not significantly) lower ACOT means on number of words and sentences. ACOT advantages that approximated significance were lower error rates in singular/plural uses and tense shifts. The only identifiable advantage for the non-ACOT group was use of more details in their writing ($p < .05$).

Discussion

Findings from the present study suggest that high computer access in classrooms creates opportunities to expand and enhance students' writing skills. Due to the complexity and naturalistic context of the experimental program examined, the results are not conclusive regarding the effectiveness of specific activities. However, interesting insights were provided regarding the strengths and weaknesses of particular writing components as well as desirable directions for the design of future programs in which computers are employed as a writing tool.

Writing Practice in a Computer-Intensive Environment

Similar to the conventional class, ACOT students received in-class instruction on writing and reading skills for approximately 45 minutes hours a day. However, in contrast to the former, they used word-processing for most home and in-class writing assignments.

With regard to integrating computer-based writing into the classroom, a highly positive outcome was the apparent ease with which the mechanics of word-processing were learned so that computer writing was increasingly used as an alternative to manual writing. Although students received initial training on the word-processor, it was brief and restricted to the fundamentals needed to begin experimenting on their own. From that point on, a "learning by doing" orientation, with supplementary group or individual instruction, was used. Much greater difficulty was encountered in implementing computer-specific (i.e., new or unique) applications, specifically electronic mail and writing-oriented CAI, that changed the conventional way that the class and its writing activities were structured. As experience with computer-intensive environments increases, so should the confidence and knowledge base needed to implement more creative and varied activities.

Writing on the BBS

There was a fairly strong consensus among all participant groups that the electronic mail activities were not a significant factor in enhancing formal writing skills. Tutors and teachers perceived the electronic mail as having substantial potential for improving writing attitudes and skills by establishing writing as a natural means of communicating with friends and tutors. Unfortunately, during this initial year of the program, writing activities were constrained by the difficulty of accessing and writing on the BBS. Many students expressed difficulty in understanding the formal writing assignments and the corrections that the tutors made.

Writing Outcomes

The clearest finding in this study emerged from reactions of ACOT students regarding their writing experiences. There was strong agreement that word-processing was enjoyable and helpful to them in completing assignments. Every ACOT student indicated that, if given a choice between a word-processor and paper-and-pencil to write an essay to be graded, they would select the former. The ability to edit easily was identified as the main advantage. Whether writing skills actually improved, however, is not as clear, but with regard to motivational-affective benefits, students were quite convincing in their support for word-processing as a desirable and preferred writing mode. Corroborative evidence for these reactions was provided in the present study by the teacher and coordinator reports, and in related research which showed ACOT students to make extensive use of the word-processor at home (Ross, Smith, Morrison, & O'Dell

1989).

Results of the writing analysis generally favored the ACOT students over the control group. Specifically, on writing Sample #1, the ACOT students wrote longer essays, used more topic sentences, and had fewer mistakes in singular/plural forms. On the Sample #2, they had fewer errors in capitalization, spelling, and subject-verb agreement, and more unity/coherence in their content. Given the inability to rule out the numerous internal validity threats associated with the present design, these results must be viewed cautiously.

Conclusions

Several conclusions were supported regarding the use and impact of computer-based writing activities in the ACOT setting. First, in replicating previous findings (Madian, 1986; Quellmaltz, 1989; Rodriques, 1985), it was found that students very much enjoyed word-processing, especially the ability to edit easily, and preferred it to the paper-and-pencil mode.

Second, given high computer access in school and at home, students employed word-processing routinely for completing writing assignments. In supporting this idea, the present study extends earlier research to contexts in which word-processing is used as a primary mode of writing rather than as an isolated activity practiced for a brief time each day or for a limited period during the school year (see review by Bangert-Drowns, 1989).

Third, this complete integration of word-processing with school work does not appear to require any special preparation for students or teachers, except with the mechanics of the word-processing program. It was interesting to observe how rapidly students learned those mechanics through independent practice and discovery, and how uninhibited they were about experimenting with different procedures compared to the college students whom the authors have taught.

Fourth, the use of electronic mail for exchanging writing messages appears to have potential as a means of developing writing skills but was weakened in the present study by hardware limitations and the absence of a structured activity plan. Activities that directly stimulate and reinforce effective writing, such as having students check each other's work and make editing suggestions (see, e.g., Bruce, et al., 1985) would seem beneficial for such purposes.

References

Baker, E., & Herman, J. (1989). *The ACOT report card: Effects of complex performance and attitude*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Bangert-Drowns, R.L. (1989). *Research on wordprocessing and writing instruction.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Becker, H.J. (1987). *The impact of computer use on children's learning: What research has shown and what it has not*. (Report No. 18) Baltimore, MD: Center for Research on Elementary and Middle Schools, The Johns Hopkins University.

Bruce, B., Michaels, S., & Watson-Gegeo, K., (1985). How computers can change the writing process. *Language Arts*, *67*(2), 143-149.

Calkins, L. (1986). *The art of teaching writing*. Portsmouth, NH: Heinemann Educational Books.

Campbell, D.J., & Stanley, J.C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Clark, R.E. (1983). Reconsidering research on learning from media. *Review of Educational Research*, *53*, 445-459.

Collins, A., Bruce, B., & Rubin, A. (1982, February). Microcomputer-based writing activities for the upper elementary grades. *Proceedings of the Fourth*

International Congress and Exposition of the Society for Applied Learning and Technology, Orlando, FL.

Concialdi, L. (1985). Is computer literacy/keyboarding elementary, Dr. Watson? *Balance Sheet*, *66*, 21-23.

Daiute, C. (1985a). Issues in using computers to socialize the writing process. *Educational Communication and Technology Journal*, *33*, 41-40.

Daiute, C. (1985b). *Writing and computers*. Reading, MA: Addison-Wesley Publishing Company, Inc.

Frankeberger, L. (1985). Keyboarding: Success in elementary grades. *Balance Sheet*, *66*(3), 41-43.

Heebner, A. (1988). Integrating word-processing with writing instruction: A review of research and practice. *Special Interest Group on Computer Uses in Education* (SIGUE), *20*, 55-64.

Hester, J. (1988). *An analysis of the alphanumeric keyboarding skills taught and achieved in the Apple Classroom of Tomorrow Program*. Unpublished Master's Thesis, Memphis State University.

Kitabchi, G. (1987). Final report for the evaluation of the Apple Classroom of Tomorrow Project Phase II. Memphis, TN: Memphis City Schools.

Kitabchi, G. (1988, June). Final report for the evaluation of the Apple Classroom of Tomorrow: Project phase III. Memphis, TN: Memphis City Schools Division of Research Services.

MacArthur, C.A. (1988). The impact of computers on the writing process. Exceptional Children, 54, 536-542.

MacGregor, S.K. (1986). Computer-assisted writing environments for elementary students. Paper presented at the National Educational Computing Conference, Louisiana State University.

Madian, J. (1986). New flexibility in curriculum development through word processing. Educational Leadership, 43,(6), 22-23.

Newman, J.M. (1984). Language arts. Language, Learning and Computers, 61 (5), 494-497.

Quellmatz, E. (1989). Freeing thoughtfulness, flexibility, and fluency: Assessment of writing in a third-grade ACOT classroom. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Rodrigues, D. (1985). Computers and basic writers. College Composition and Communication, 36, 336-339.

Ross, S.M., Morrison, G.R., Smith, L.J., & Cleveland, E. (1989). Alternative programs for distance tutoring of at-risk elementary school children. Computers in human behavior. Manuscript submitted for publication.

Ross, S.M., Smith, L.J., Morrison, G.R., & Erickson, A. (1989). Helping at-risk children learn through distance tutoring: The Memphis ACOT program. Technological Horizons in Education, 16, 68-71.

Ross, S.M., Smith, L.J., Morrison, G.R., & O'Dell, J.K. (1989). What happens after ACOT: Outcomes for program graduates one year later. Final report for Apple Classroom of Tomorrow Research Grant. Memphis, TN: Memphis State University.

Slavin, R.E. (1986). Best evidence synthesis: An alternative to meta-analytic and traditional reviews. Educational Researcher, 15 5-11.

Slavin, R.E., & Madden, N.A. (1989). What works for students at risk: A research synthesis. Educational Leadership, 4-12.

Strictland, J. (1986, Spring). Beyond word processing: Computer-assisted writing instruction. Journal of Teaching Writing, 51-58.

Sudol, R.A. (1985). Applied word processing: Notes on authority, responsibility, and revision in a workshop model. College, Composition, and Communications, 36(3), 331-335.

Wetzel, K.A. (1985). The effect of using the computer in a process writing *program on improving the writing quality of third, fourth, and fifth grade pupils*. Unpublished doctoral dissertation, University of Oregon.

Table 1

Comparisons Between ACOT and Control Writing Sample #1

Variable	ACOT Mean	Control Mean	t value	Probability
No. of words	80.35	53.17	3.12	.004
No. of sentences	6.87	5.21	1.91	.063
Topic sentence[a]	.74	.38	2.64	.011
No. of details	4.91	5.00	-.14	.889
No. of misspellings	3.22	2.13	1.63	.112
No. of fragments	.83	.83	-.02	.99
No. of run-ons	.74	1.08	-1.13	.263
Capitalization error	.65	1.25	-1.50	.142
Punctuation error	3.57	3.67	- .15	.883
Pronoun score[a]	.17	.67	-3.85	.000
Articles error	.61	.42	.60	.551
Singular/plural error	.26	.88	-2.49	.018
Subject/verb error	1.13	.92	.44	.660
Tense shift error	.22	.25	- .21	.832
Redundancy score[a]	.61	.71	-.71	.483
Content[b]	2.87	2.62	1.05	.298

[a] Scores were based on a dichotomous, satisfactory (1) or unsatisfactory (0) system.

[b] Scores were based on a five-point system (5=highest).

Table 2

Comparisons Between ACOT and Control Groups on Writing Sample #2

Variable	ACOT Mean	Control Mean	t-value	Probability
No. of words	91.63	115.70	-1.26	.216
No. of sentences	5.79	7.50	-1.18	.247
Topic sentence[a]	.88	.65	1.74	.091
No. of details	4.08	6.10	-2.12	.045
Correct letter form[a]	.42	.55	- .87	.391
No. of misspellings	2.83	7.05	-2.55	.019
No. of fragments	.88	1.30	- .80	.432
No. of run-ons	.54	1.00	-1.64	.110
Capitalization error	.38	3.30	-2.19	.041
Punctuation error	2.83	5.15	-2.02	.054
Pronoun score[a]	.83	.90	- .64	.524
Article error	.33	.80	-1.36	.186
Singular/Plural error	.67	1.60	-1.73	.096
Subject/verb error	.13	.70	-2.64	.015
Tense shift error	.21	.55	-1.95	.061
Redundancy score[a]	.75	.60	.93	.359
Content[b]	2.96	2.35	2.40	.021

[a] Scores were based on a dichotomous, satisfactory (1) or unsatisfactory (0) system.

[b] Scores were based on a five-point system (5=highest).

Author's Note

This research was supported by Apple Classrooms of Tomorrow. A longer version of this paper has been accepted for publication in Journal of Computing in Childhood Education.

Title:

Receive-site Facilitator Practices and Student Performance in Satellite-delivered Instruction

Author:

Floyd K. Russell

RECEIVE-SITE FACILITATOR KNOWLEDGE AND PRACTICES AND STUDENT ACHIEVEMENT IN A SATELLITE COURSE FOR HIGH SCHOOL STUDENTS

FLOYD K. RUSSELL, JR.

ABSTRACT

This study examined the relationship of satellite distance education receive-site facilitator knowledge and practices to student achievement in a satellite distance education course. The specific areas of interest were: 1) facilitator knowledge about the satellite telecommunications delivery system, as evidenced by a knowledge test and a self-report performance inventory; and, 2) facilitator practices for enhancing student interaction with the distance teacher and the course content, as measured by a self-report performance inventory of facilitator practices. Students' semester grades were used as the measure of student achievement which was compared to facilitator knowledge and practices.

An unbalanced 2 x 2 factorial design was used to analyze data for the study questions. Both independent facilitator variables, knowledge of the satellite telecommunications delivery system and practices, were assigned to high and low groups around a median. Scores above the median were high scores; scores below the median were low scores. The dependent variable was the site mean for student's final semester grade.

Analysis of the study questions did not reveal a significant relationship between student achievement and either the facilitators' knowledge of the satellite system or practices for facilitating student interaction with the course content and the distance teacher or their interaction.

A relationship may not exist between facilitator knowledge and practices attributes and student achievement. Alternatively, the instruments used with the population in this study may have failed to find a relationship. It is possible that a test with different items may find a relationship between facilitator attributes and student achievement.

It is possible also that a test alone may not reveal a relationship between facilitator attributes and student achievement. This may, in fact, be the case because the data indicated that interacting factors are the best indicators of student achievement across sites. Those facilitators best able to assist students in a variety of situations and resolve problems of both an instructional nature and a system operations nature may be more successful in facilitating student achievement. That type of facilitator would be more likely to fit a model of a facilitator capable of reducing "structure" in a system and increasing "dialog".

RECEIVE-SITE FACILITATOR PRACTICES AND STUDENT PERFORMANCE IN SATELLITE-DELIVERED INSTRUCTION

Introduction

This study examined the relationship of selected attributes of receive-site facilitators for satellite-delivered distance education to student achievement in a distance education course. The specific facilitator attributes of interest were facilitator knowledge of the satellite telecommunications delivery system and facilitator practices for assisting students with satellite-delivered instruction.

While satellite communications technology has proven to be a viable means of providing courses which otherwise might not be offered to secondary education students, the complexity of the enterprise has had an impact on how education is organized and managed. When classes have been taught via satellite by a distance teacher, a classroom facilitator at the receive-site has worked with students to help them interact with the distance teacher and the course content. The facilitator has acted in a support role. The distance teacher has had primary responsibility for student learning.

Despite prior experiences with satellite delivered instruction in the secondary education environment in the United States (Grayson, 1974; Vest, 1975), there is not a large body of research available on the effectiveness or patterns of use. This is due in part to the relatively short period of time for intensive use of the technology in other than demonstration projects. The 1988 Office of Technology Assessment (OTA, 1988) report, Power On! New Tools for Teaching and Learning, states: "Little research exists that specifically addresses K-12 distance education, and what does exist is limited in scope and often anecdotal (p. 45)." Other researchers (Eiserman & Williams, 1987) have found it necessary to draw on research at the higher education level in order to develop information on effectiveness and practice of distance education.

The Office of Technology Assessment's 1989 report, Linking for Learning (OTA, 1989, p. 97), cites summary findings of research that found that students were "consistently more successful" in distance education classes where facilitators encouraged students and learned and participated with them in the course.

Facilitator's Role

While receive-site facilitators for satellite distance education are not the primary teacher, they are in the classroom with the students, and they fulfill the face-to-face contact role that the distance teacher cannot always fulfill due to time or distance. This team relationship between facilitator and distance teacher makes it necessary to give consideration to how the facilitator uses technology at the receive-site to help accomplish the instructional objectives of the distance teacher and the organization providing the course. Facilitators have two types of opportunities to impact student achievement. One type of opportunity is that of serving as a role model for the use of technology as a tool for learning. The other opportunity is that of motivating the student to succeed and monitoring student activity and interaction with the course content and the distance teacher.

Regular exchanges of information and a team approach are emphasized by distance education course providers. Facilitator responsibilities (TI-IN, 1989, p. 7; OTA, 1989, p.97), include:

1. Attending training sessions
2. Monitoring equipment operation
    a. Demonstrating proper use of equipment
    b. Reporting technical difficulties
    c. Recording courses for replay
3. Supporting instruction
    a. Disseminating and returning course materials
    b. Monitoring and facilitating instruction
    c. Keeping records
    d. Consulting with the distance teacher
    e. Administering evaluations
    f. Assisting with educational activities as assigned by the distance teacher

Facilitator's Importance to Learning

It is important to know what impact facilitators may have on student learning in a distance education course. Three considerations are of central importance:

1. The distance education system must operate successfully if learning is to take place (technical operation).
2. The facilitator acts to focus the student's attention on the distance teacher and the course content. The smoother the operation, the less interference in learning (utilization methods).
3. The technology, itself, may serve as an organizer for student learning. Smoother operation contributes to better organization of content by the students (combination of technical operation and utilization methods).

Methods

The research undertaken in this study was designed to develop and pilot the use of a model suitable for examining the effectiveness of various aspects of a distance education delivery system. The receive-site facilitator was chosen as the aspect or system component to study.

Study Questions

The relationship of specific facilitator knowledge and practices to student achievement was examined by asking these questions:

Study Question 1

What is the relationship of a receive-site facilitator's knowledge of the satellite telecommunications delivery system, as indicated by a knowledge test and self-report inventory, to student achievement, as measured by the final semester grade (numerical range of 68-99), for a satellite delivered course?

Study Question 2

What is the relationship of a receive-site distance education facilitator's practices for enhancing student interaction with the distance teacher and the course content in satellite delivered instruction, as indicated by a self-report inventory of practices, and student achievement, as measured by the final semester grade, for a satellite delivered course?

Study Question 3

Is there an interaction between a receive-site facilitator's knowledge of the satellite telecommunications delivery system and the facilitator's practices for enhancing student interaction with the distance teacher and the course content in satellite delivered instruction which relates to student achievement in a satellite delivered course?

Research Model

In proposing possible research areas for the use of new technologies in education, Clark and Salomon (1986, p. 474) suggest that researchers avoid looking for differences in how well a technology enables teaching. They state: "Any new technology is likely to teach better than its predecessor because it generally provides better prepared instructional materials and its novelty engages learners (p. 474)."

Clark and Salomon suggest inquiry into a medium's impact on the environment into which it is introduced. Inquiries into the impact of a technological medium of instruction may address changes in the way resources are allocated, changes in how providers and users interact, and changes in the structure of organizations. Two aspects of this type of problem are of interest: 1) how innovations naturally influence the educational setting and its attendant resources , and 2) how these innovations can be directed in order to achieve desirable outcomes.

Coldeway (1988) stated that "The components that directly impact learners are of particular interest to many involved in distance education, especially components that may impact learning, achievement, motivation, rates of study, errors, communications, etc. (p. 50)." First stage component research, which analyzes a component for existence and level of quality, is a prerequisite for understanding how the components of a system may affect learning; subsequent research may manipulate the component in order to study changes in the effect (Coldeway, 1988, p. 51).

[PLEASE PLACE TABLE 1 ABOUT HERE]

Transactional distance, or the relative relationship of "dialog" and "structure" in an educational endeavor, can be used as a measure of the degree of separation between teacher and learner. Moore (1983, p. 157) defines dialogue as "...the extent to which, in any educational programme [sic], learner and educator are able to respond to each other." Educational programs can be judged more or less distant on the basis of the level of dialogue [sic] between learner and teacher. Moore defines structure as "...a measure of an educational programme's [sic] responsiveness to learners' individual needs." More distant programs have a higher structure and are less responsive to an individual learner's needs. This relationship is shown in Figure 1.

[PLEASE PLACE FIGURE 1 ABOUT HERE]

Content, educational philosophy, learner and educational personalities, the learning environment, and the communications medium determine the nature of the dialog. Structure includes educational objectives, teaching strategies, and evaluation methods. Flexibility in the preparation and application of objectives, strategies, and evaluation indicates less structure (Moore, 1983, p. 157).

The use of media in distance education provides a means of changing the level of dialogue and structure. Media can be used to optimize interaction between the learner and the teacher by providing a balance between dialogue and structure (Saba, 1988, p. 17).

An organizational model for satellite delivery systems in education, proposed by Vest (1975), has provided a framework for studying the facilitator's role in secondary education. In this system, the facilitator can be viewed as part of the classroom subsystem in the larger satellite telecommunications delivery system for secondary level distance education. Vest (1975 pp. 253-266) described a system for organizing and coordinating a complex satellite educational delivery system. The system contained nine subsystems. The subsystems and the relationships among them are shown in Figure 2. A facilitator for a distance education class may be considered to be a component of the classroom subsystem in Vest's model.

[PLEASE PLACE FIGURE 2 ABOUT HERE]

The model for this study, then, is first stage component research which inquires about the impact of a component in a subsystem of the larger system of satellite-delivered instruction by examining the existence and level of quality of what receive-site facilitators do to assist students with distance education and its relationship to student achievement.

Design

An unbalanced 2 x 2 factorial design was used to analyze data for the Study Questions. The design provided for analyzing four different combinations of facilitator attributes: 1) High Practices/High Knowledge (HH), 2) High Practices/Low Knowledge (HL), 3) Low Practices/High Knowledge (LH), and 4) Low Practices/Low Knowledge (LL). Total N for facilitators was ten, but the Ns for the groups were unbalanced; and, therefore, a regression model was used for the analysis.

Both independent facilitator variables, facilitator knowledge of the satellite telecommunications delivery system and facilitator practices, were assigned to high and low groups around a median. Scores above the median are high scores; scores below the median are low scores. The dependent variable was the site mean for student's final semester grade. The unbalanced 2 x 2 factorial design is illustrated in Figure 3.

[PLEASE PLACE FIGURE 3 ABOUT HERE]

Subject Selection

Students enrolled in a satellite distance education French language course during the Fall Semester, 1989, and their facilitators were asked to participate in the study. Twenty-seven receive-site classes, with 111 students, were available for the study. In order for a site to be included in the study, the facilitator for the site had to agree to be included in the study. Thirty-four students (31 %) from ten sites and their facilitators (ten; 37%) responded to the invitation to participate in the study.

Typically, students enrolled in distance education courses are screened by guidance counselors or principals for ability to do advanced work. Of the thirty-four students (31%) who responded to this survey, 13 (38%) were male and 21 (62%) were female. The average student respondent for this French I distance education course was 16 years of age, in a school grade with 84 students, and enrolled in a receive-site class of seven students.

The facilitators for the distance education course used in this study were either teachers appointed to be facilitators or aides hired to be facilitators. The facilitators in the study were not required to be certified in the subject area of the satellite course.

Contact with Subjects

Researcher contact with both students and facilitators took place through written communication which accompanied the data collection instruments. Data collection instruments for both facilitators and students were distributed by the satellite system administrator through their normal administrative channels for the French language course. All data instruments were collected as a package at each receive-site and returned to the researcher through the satellite system administrative channels. The researcher was available by telephone to answer questions from facilitators, students, parents or guardians, or the satellite system administrator; however, no contact, except for administrative liaison through satellite system personnel, was made with the researcher by subjects or the distance teacher.

Both facilitator and student background data were collected at the beginning of the Spring (2nd) Semester 1990. Student data collection was performed by the receive-site facilitator in conjunction with the satellite system administration. Facilitator data collection was performed by the satellite system administration through administrative channels used for communication with site facilitators. These channels included surface mail, telephone calls, and announcements by the distance teacher over the satellite television system.

Collection of Student Data

A survey questionnaire form was used to collect student background data. The researcher-designed student background survey was reviewed for consistency and readability by a panel consisting of teachers and administrators familiar with the expected age group range of the students. The panel included secondary level distance education site facilitators, a state department of education administrator with responsibility for distance education, a high school level writing teacher, and a higher education reading researcher. The survey instrument used in the study incorporates suggestions made by the review panel for improving consistency, readability, and relevancy of the instrument for the student population available to the researcher.

Student achievement in the course was determined by assignment of a final grade for the semester (Fall Semester 1989) preceding data collection. The final semester grade for students was given by the course provider in accordance with the normal procedure they observe for testing students in a distance education course. Assigned grades ranged from 68 to 99. The semester ran from August 30, 1989 to January 12, 1990. Semester grades (68 to 99) were based on equal weights for each of the six weeks grading periods and the semester examination. Six weeks grades were based on major tests and homework assignments with major tests receiving more weight. The student's first six weeks grade, final examination score, and semester grade were obtained from satellite system administrative records.

Collection of Facilitator Data

Facilitator data collection instruments included: 1) a background survey form and 2) a knowledge test and self-report inventory for the satellite telecommunications delivery system, and 3) a self-report inventory of practices for facilitating satellite courses.

Facilitator data collection instruments were researcher-designed based on lists of duties for site facilitators, issues identified in the review of literature for the study, and discussions with researchers and practitioners in distance education. The instruments were reviewed by means of a modified one-round Delphi procedure by a panel consisting of an industry training manager, higher education researchers, K-12 and state department level administrators, and secondary level distance education site facilitators. Comments and suggestions for revision were closely aligned, and one round was determined to be sufficient for the instrument review.

Data analyses for the study included:

For facilitator data (N=10):

1. Relationship of facilitator knowledge, as indicated by the score on the satellite telecommunications delivery system test, to student group achievement in the course, as measured by the final course grade.
2. Relationship of facilitator practices, as indicated by the score on the practices inventory, to student group achievement in the course, as measured by the final course grade.
3. Interaction of facilitator knowledge and practices with student group achievement, as measured by the final course grade.

For student data (N=34):

4. Differences of final grades among student groups across sites.
5. Description of which variables, including facilitator qualifications, are most predictive of final grade for students.

Satellite Telecommunications Delivery System Knowledge Test

The knowledge test was an assessment of the facilitator's knowledge about management and operation of the satellite telecommunications delivery system. Questions on this test measured facilitator knowledge in two areas: 1) general knowledge about satellite systems and 2) general operation knowledge about classroom receive-site equipment.

Facilitator Practices Self-Report Inventory

This inventory measured facilitator practices related to enhancement of the students' interaction with the distance teacher and the course content. The inventory was constructed from a list of duties for facilitators. It includes knowledge items and frequency of occurrence items.

Results

Data analysis was accomplished with the aid of FASTAT™: Fast Statistics for the Macintosh (Ver. 1) statistical analysis software running on a Macintosh IIx microcomputer and SYSTAT©: The

System for Statistics (Ver. 3.0) software running on an IBM PS/2 Model 50 microcomputer.

Findings Related to Study Questions

Neither Knowledge Category nor Practices Category nor their interaction had a significant relationship with Semester Grade which had a mean of 88.206 and a standard deviation of 7.804. The combination of Knowledge Category and Practices Category and their interaction accounts for only 9.7% of the variance in semester grades across sites as indicated by the value for Squared Multiple R for this regression. The data indicate that the probability of the Knowledge Category contribution to Semester Grade being due to chance could be .865 and that the probability of the Practices Category contribution to the relationship being due to chance could be .983. This analysis is located in Table 2.

[PLEASE PLACE TABLE 2 ABOUT HERE]

Subsidiary Findings

The Stepwise regression procedure in SYSTAT (Ver 3.0) was used to develop a model for predicting student achievement. The procedure started with factors with a high correlation with Semester Grade, and it produced the model in Table 3. This model accounted for 87.1% of the variance in Semester Grade ($F=15.767$, $p \leq .000$). The data indicated that the combination of Six Weeks Grade (ability indicator), Availability of Printed Handouts(scale = 1-5), and Facilitator Encouragement(scale = 1-5),and Site (each contrasted to Site 1) had a relationship with Semester Grade.

When the first six weeks grade is used a covariate with semester grade, the combination of the availability of print handouts when they are being discussed, how much the students feel the facilitator encourages activities which help with learning material presented in the course, and site has a significant relationship ($a=.05$) with semester grade .

The problem with this model is that Site, which is an undefined factor) is one of the factors and that Print Handout has a negative influence. Intuitively, learning aids would be expected to enhance learning. The data regarding availability of printed handouts in this study is confounded because of conflicting reports concerning this factor. It is also possible that the large percentage of sites missing handouts has resulted in homogeneous data which does not have enough variance to allow a finding of the real impact of printed handouts. Site differences could not be accounted for with the data available.

[PLEASE PLACE TABLE 3 ABOUT HERE]

This model uses the six weeks grade as a factor in student ability, the degree of facilitator encouragement, the availability of handouts, and the differences in sites to account for 87.1% of the variance in Semester Grade across sites. Sites 2 through 7 are different from Site 1. Inspection of the data did not reveal reasons for the similarities in contribution to Semester Grade for students being at Sites 3, 4, and 5 or the similarities in contribution to Semester Grade for being at Sites 6 and 7.

Implications

This study was limited to a French language satellite-delivered distance education course in secondary schools. The model for the study was limited to one course provided through a single satellite system. The study assumed that the learner was not in direct control of the delivery system and that the distance teacher was not in direct, continuous interaction with the student. The study also assumed that the contribution of the distance teacher, who has primary responsibility for student learning, to student achievement balanced out across the study.

The analyses for relationship of facilitator knowledge and practices with student achievement did not reveal a significant relationship between student achievement for either the knowledge category or the practices category or their interaction. It is possible that what facilitators know about the satellite delivery system and their practices for enhancing student interaction with the distance teacher and the course content do not have a relationship with student achievement. Alternatively, a relationship, which the instruments used with the population in this study failed to find, may exist. It is possible that a different test with different items may find a relationship between facilitator attributes and student achievement.

It is possible also that a test alone may not reveal a relationship between facilitator attributes and student achievement. This may, in fact, be the case because the data indicated (Table 3) that a combination of factors are the best indicators of student achievement and differences in sites.

Recommendations

There are many possibilities for further research into facilitators' relationships to student achievement. This study could be repeated with a larger sample. Because this study had a relatively small sample of students and facilitators to work with, there is the possibility that the data were homogeneous and did not have differences that were detectable.

Other researchers may want to consider replicating the study across content areas and across course providers. In either or both cases, further studies should consider employing a means of comparing student achievement to standardized norms in order to determine the relative effectiveness of the distance education programs.

Recommendations for improving the data collection are: to include measures of the learning styles and aptitudes of students, to include more measures of things effective teachers do, and to include qualitative observation of students and facilitators. It also may be useful to include some measures of the difference in recall of information presented through audio and video modes. Researchers conducting research similar to this study may wish to collect data in separate categories for when students view video recordings for review and when they view recordings to make up lessons. Separating these two categories and asking if printed handouts were available when the distance teacher talked about the material may eliminate the confusion about the importance of printed handouts encountered by this study.

Qualitative research may be the best methodology for determining what facilitators and students actually do in distance classes. This type of research may provide valuable insight into what facilitators know about facilitating and how they facilitate student interaction with the distance teacher and the course content. It may also provide information about other factors which influence student achievement.

Follow-up studies that track students' ability to do college-level work in the subject areas in which they took distance courses may provide information about the effectiveness of distance education as a system for providing pre-college and advanced placement courses.

Further research of the type conducted in this study may be most useful if it can point to better ways that the course provider, distance teacher, and facilitators can work together as a team. Due to the complexity of conducting research in a consortium environment, researchers may wish to consider contractual arrangements for working with course providers and satellite system administrators to conduct research into delivery of instruction by satellite.

References

Clark, R. E., & Salomon, G. (1986). Media in teaching. In M. C. Wittrock (Ed.), Handbook of research on teaching (3rd. ed.) (pp. 464-478). New York: MacMillian.

Coldeway, D. O. (1988). Methodological issues in distance education research, The American Journal of Distance Education Research, 2(3), 45-54.

Eiserman, W. D. & Williams, D. D. (1987). Statewide evaluation report on productivity project studies related to improved use of technology to extend educational programs. Sub-report two: Distance education in elementary and secondary schools. A review of literature. Logan, UT: Wasatch Institute for Research and Evaluation [for Utah State office of Education; presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 23-27, 1984]. (ERIC Document Reproduction Service No. ED 291 350)

Grayson, L. P. (1974). Educational satellites: The ATS-6 experiments. Journal of Educational Technology Systems, 3(2), 89-123.

Holmberg, B. (1986). Growth and structure of distance education. London: Croom Helm.

Moore, M. (1983). The individual adult learner. In M. Tight (Ed.), Education for Adults: Vol. 1. Adult Learning and Education, pp. 153-168. London: Croom Helm.

Saba, F. (1988). Integrated telecommunications systems and instructional transaction. The American Journal of Distance Education, 2(3), 17-24.

TI-IN Network. (1989) TI-IN United STAR Network Operational Guidelines for Direct Student Instruction and Staff Development. San Antonio, TX: TI-IN Network.

U. S. Congress, Office of Technology Assessment. (1988). Power on! New tools for teaching and learning (OTA Publication No. OTA-SET-379). Washington, DC: U. S. Government Printing Office.

U. S. Congress, Office of Technology Assessment. (1989). Linking for learning: A new course for education (OTA Publication No. OTA-SET-430). Washington, DC: U. S. Government Printing Office.

Vest, C. R. (1975). The educational satellite: A systems model. Journal of Educational Technology Systems, 3(4), 249-266.

Tables

Table 1: Conceptual Framework for Distance Education Research

Input Variables	Process Variables	Outcome Variables
Student •Educational background •Perceived needs •Motivation •Learning styles •Study environment	**Development** •Curriculum development model •Design of instruction •Media •Course workload •Pacing •Production procedures	**Student** •Enrollment •Academic progress •Academic performance •Use of materials and services •Dropout
System •National requirements •Institutional policy •Financial resources •Technological resources •Human resources •Geography	**Delivery** •Recruitment methods •Academic support •Formal feedback	**System** •Development efficiency •Cost •Effectiveness •Acceptance in the system

Note. From "Methodological Issues in Distance Education Research" by D. O. Coldeway, 1988, *The American Journal of Distance Education, 2*, p. 52. Copyright 1988 by *The American Journal of Distance Education.* Adapted by permission.

Table 2: Regression of Semester Grade on Knowledge Category and Practices Category

Dep var: SEM_GD N=10 Multiple R : 311 Squared Multiple R: .097
Adjusted Squared Multiple R .000 Standard Error of Estimate: 8.666

Variable	Coefficient	Std. Error	Std. Coef	Tolerance	T	P (2 tail)
Constant	83.540	37.773	0.000	.	2.212	0.069
KNOW_CAT	3.84	21.664	0.271	0.0640000	0.177	0.865
PRAC_CAT	-0.478	21.664	-0.034	0.0640000	-0.022	0.983
KNOW_CAT x PRAC_CAT	0.437	13.702	0.043	0.0816327	0.032	0.976

Analysis of Variance

Source	Sum-of-Squares	DF	Mean-Square	F-Ratio	P
Regression	48.162	3	16.054	0.214	0.883
Residual	450.564	6	75.094		

Case D is an outlier. (Studentized residual = -3.352

Table 3: Regression of Semester Grade on Six Weeks Grade, Site Number, Print Hand Out Availability, and Facilitator Encouragement

Dep var: SEM_GD N=31 Multiple R : 8933 Squared Multiple R: .871
Adjusted Squared Multiple R .816 Standard Error of Estimate: 3.409

Variable	**Coefficient**	**Std. Error**	**Std. Coef**	**Tolerance**	**T**	**P (2 tail)**
Constant	-0.364	10.101	0.000	1.0000000	-0.036	0.972
Sx. W. Gd.	0.589	0.121	0.521	.5375250	4.877	0.000*
Prt. HO	-3.857	1.529	-0.893	.0490236	-2.523	0.020*
Fac. Enco.	4.965	1.563	0.390	.4072758	3.176	0.005*
SITE2	15.059	2.694	0.875	.2507826	5.590	0.000*
SITE3	26.128	6.428	1.321	.0581407	4.065	0.001*
SITE4	27.783	6.133	1.404	.0638530	4.530	0.000*
SITE5	25.647	4.717	0.806	.2792216	5.437	0.000*
SITE6	34.608	6.384	1.485	.0818683	5.421	0.000*
SITE7	33.735	6.916	1.060	.1298824	4.878	0.000*

Analysis of Variance

Source	**Sum-of-Squares**	**DF**	**Mean-Square**	**F-Ratio**	**P**
Regression	1649.333	9	183.259	15.767	0.000*
Residual	244.087	21	11.623		

*p≤.05

Figures

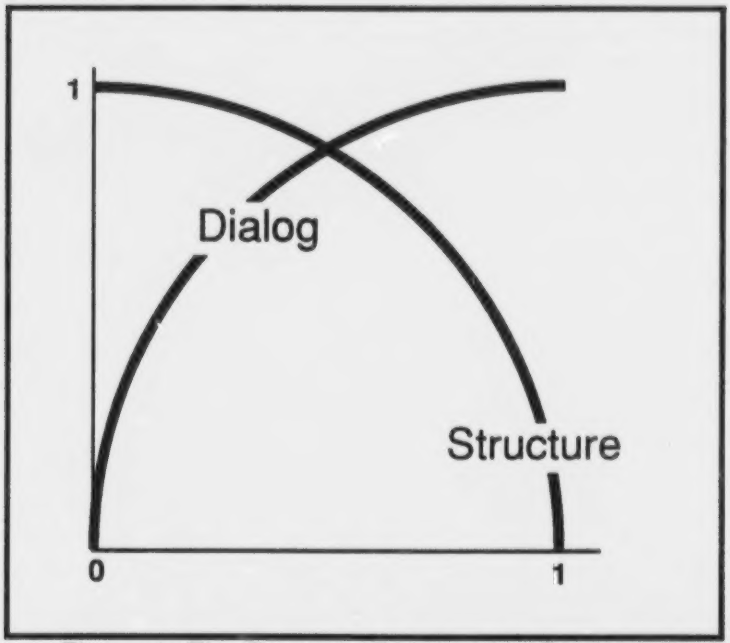


Figure 1. The Dynamic Relationship Between Dialog and Structure

Note. From "Integrated telecommunications Systems and Instructional Transaction" by F. Saba, 1988, The American Journal of Distance Education, 2, p. 23. Copyright 1982 by The American Journal of Distance Education. Used by permission.

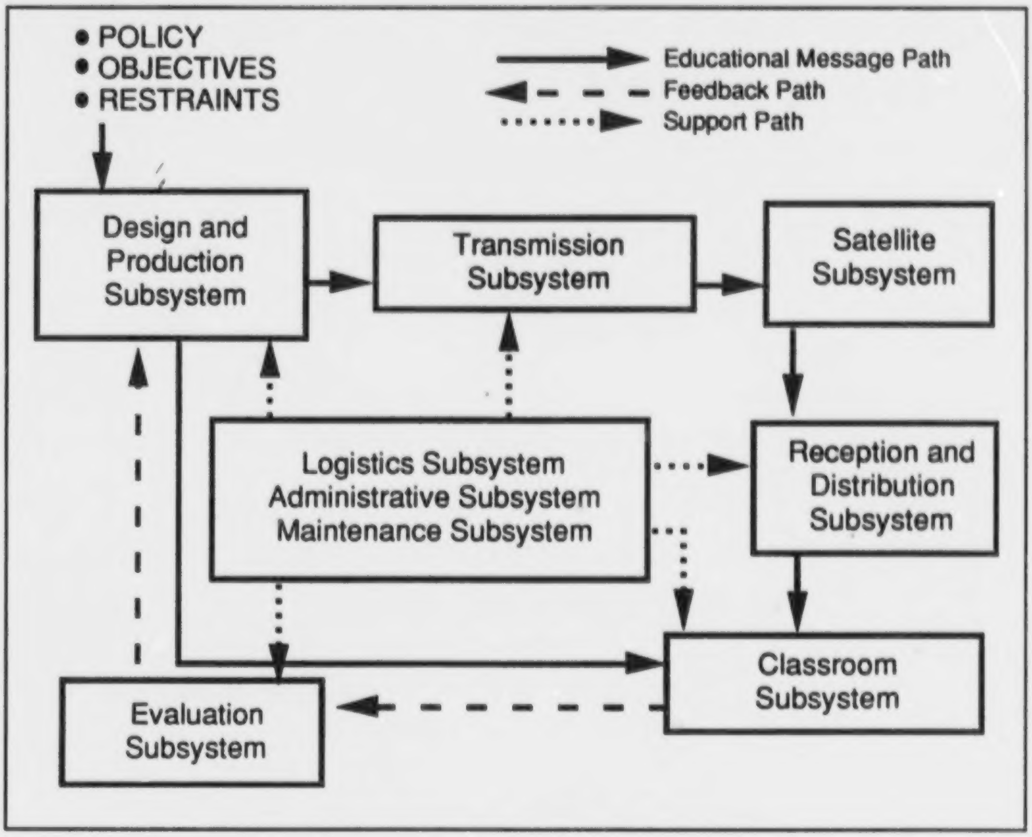

• POLICY
• OBJECTIVES
• RESTRAINTS
Educational Message Path
Feedback Path
Support Path
Design and Production Subsystem
Transmission Subsystem
Satellite Subsystem
Logistics Subsystem
Administrative Subsystem
Maintenance Subsystem
Reception and Distribution Subsystem
Classroom Subsystem
Evaluation Subsystem

Figure 2. The Educational Satellite Systems Model. Note. From "The education Satellite: A Systems Model" by C. R. Vest, 1975, Journal of Educational Technology Systems, 3, p. 254. Copyright 1975 by Baywood Publishing Co. Adapted by permission.

	Knowledge	
Practices	High	Low
High	HH	HL
Low	LH	LL

Figure 3. Design for Analysis of Variance for for Unequal Groups: Receive-Site Facilitator Knowledge and Practices Relationship to Student Achievement

Title:

The Use of Instructional Development Procedures to Create Exhibits: A Survey of Major American Museums

Authors:

George W. Schlenk
Sharon A. Shrock

The Use of Instructional Development Proceduresto Create Exhibits: A Survey of Major American Museums

George W. Schlenk
Sharon A. Shrock

Southern Illinois University
Carbondale, Illinois

Introduction and Problem Statement

Purpose of the Study

The purpose of the study was to investigate whether major American museums are using Instructional Development procedures to develop their exhibits. It was hoped that an accurate picture of the exhibit development process at major American museums would also be produced.

Background Information

Instructional development is the discipline which develops instructional systems, as a museum exhibit might be classified, through the use of a variety of systematic procedures. These procedures include: needs assessments, goal setting, written instructional objectives, trial runs, evaluation, and feedback utilization.

Two themes are clear from the literature of the museum field. These writers indicate that eduction is one of the museum's main aims and objectives. They also maintain that the presentation of exhibits is one of their primary areas of activity. Thus the general public would be justified in assuming that museum exhibits should be educational/instructional.

The question that arises for the researcher is whether museums are engaging in the necessary procedures to insure that their exhibits are instructional. If one compares what the instructional developer believes produces good instruction with what museum professionals are calling for to improve museums, there would be two very similar lists. The conclusion which might be drawn from this circumstance is that museum exhibits are not instructional, and further, that museum professionals recognize this situation. However, since there are also calls for better or more extensive exhibit evaluation, a more likely conclusion would be that museum professionals do not know if their exhibits are instructionally effective.

The calls by museum professionals for the use of many of the individual procedures employed by the instructional development process differ from the practice of instructional development in at least one very important way. The application of one or several of the various individual principles/procedures used in instructional development, is a piecemeal, symptomatic approach, as compared to the systematic approach of instructional development. In the past there have been calls for the use of instructional development to develop museum exhibits, most notably by Chandler G. Screven (1974). The purpose of this study was to learn if museums have followed his advice and adopted the use of instructional development to create exhibits.

Research Question

The major research question which the study sought to answer is: Are major American museums using the instructional development process to develop their exhibits? In addition to this question, the study sought to answer several secondary questions: What procedures are museums currently using to develop exhibits? What are the reasons that museums are/are not using instructional development to develop their exhibits? What has been the experience of those museums using instructional development? Do museums feel there is a need to reorganize their exhibit development processes? Would museums be open to the idea of an instructional developer/developers coordinating their exhibit development processes? What would be the reasons museum professionals might resist a change to using instructional development to develop museum exhibits?

Significance of the Problem

The museums of the United States, and indeed, the museums of the world are the storehouses of a vast collection of artifacts and information. It is these collections which document, illustrate, and even inspire much of what man knows about the earth and himself. There is much knowledge yet to be drawn from these sources, and much of what will need to be learned anew by future generations will depend on the maintenance of these collections. Thus it is imperative that these collections be preserved and utilized to their fullest by society.

Obviously, these collections have much that they can teach us. However, this cliched use of the word "teach" is really a misnomer. Museum collections can not teach us anything. What may be learned is what individuals or scholars endeavor to draw out of these objects through study and research. This process of research may entail years to divine the meaning of the artifacts which make up the collections and exhibits of our museums.

The general public, which museums endeavor to educate, does not have hours and hours of research time to study museum exhibits -- to draw out their meaning, to discover the exhibit's relevance to themselves, society, or the physical world. They should not have to rediscover that which scholars have already explored. That is not to say that the museum visitor has no responsibility to make an effort to learn or that they do not need certain prerequisite knowledge to fully appreciate the museum experience.

This does mean though, that museum professionals have a responsibility to present exhibits which are developed and designed in such a manner that everyone, from the casual visitor to the scholar doing research, can learn the insights and knowledge of museum professionals. Thus, the exhibits must be developed and presented in a systematic manner, to insure that they are instructionally sound. The principles and procedures of instructional development offer a proven systematic approach that would be ideal to develop and present exhibits.

The lineage of instructional development can be traced back to World War II, when the military was faced with the tremendous task of educating millions of people in the most efficient manner. To this end, the procedures of instructional development were first developed and introduced. Since that time, these procedures have been further researched and refined, and have continued to be used by the government and the military and are rapidly being adopted by American corporations -- all faced with the similar task of educating or training large numbers of people with diverse backgrounds to learn a large variety of skills, procedures, and basic knowledge.

In the past twenty-five years there has been a huge increase in museum attendance. This large increase and the museums' mandate to serve the general public, as well as professional scholars, would indicate that museums should utilize instructional development procedures to develop exhibits to insure that the exhibits are instructionally sound. Thus, it would seem very pertinent to question whether our museums are using instructional development to educate, that is, availing themselves of the most recent research in instructional methods. It would seem of interest whether or not our publicly supported museums are using the most effective means to allow their general audiences, as well as scholars, to learn as much as possible from their exhibits.

Since a number of museum professionals are calling for the adoption and use of various individual procedures that make up instructional development, it would be reasonable to conclude that museums are not carrying out the amalgamation of these procedures -- instructional development -- if they are not practicing the individual procedures. Thus, if it is learned where museums stand in regard to instructional development's use/non-use, or what their attitudes are toward its adoption, it might be possible to advise them on what steps to take to begin to adopt instructional development for use in their exhibit development. In this way, museums might take advantage of the same educational technology, or instructional development, to accomplish their stated educational goals, that is being successfully utilized by the United States government, the military, and corporate America.

Abbreviated Literature Review

Since the 1960s a number of museums have been aware of the need to interact and communicate with their constituency in a more overt manner to facilitate learning. Many museums tried a number of innovative programs -- neighborhood storefront museums, school programs and other outreach programs -- in addition to their exhibits. The Belmont Report (1969) lauded a number of these programs. However, some of these programs were cut back over the years due to funding problems. And while some of these exhibits and programs were successful, some were not. They frequently contained the same strengths and weaknesses as did the exhibits and programs they had offered "in house" all along.

Museums had come a long way since the days when glass cases were stuffed with objects with Latin names. Museum professionals were trying to communicate more than a Latin name -- they were trying to communicate to the visitors the importance and meaning of the objects displayed. However, there were some who wanted to limit this "interpretation" of museum collections to museum visitors. Cameron (1968) offered his view of a museum as a communication system in which museums should strictly limit their "translation" through media of the artifacts in exhibits. He believed that these artifacts comprised their own language and that visitors should learn this language.

During the 1970s, a number of people in the museum profession were struggling with how to correct or improve the weaknesses they saw in museum exhibits and educational programs -- but they had to contend with viewpoints such as that expressed by Cameron. They also had few places to turn to for help. In a number of instances they were successful, but these successes were somewhat localized -- there was no real means to learn from each other's successes and failures.

Shettel (1973) and Screven (1974), writing on instructional design and exhibits, both offered museums the systematic means to begin dealing with improving exhibits and educational programs. It was during this period that museums began looking to other fields for solutions to their problems. Miles (1982) recounts looking for these solutions and gradually finding them over a ten year period in fields like psychology, communication sciences, and educational technology. The museum literature at this time contained numerous references to museum professionals applying various techniques from the field of instructional technology and instructional design. They talked of learner/audience analysis, needs assessments, objectives, and formative evaluation. Griggs (1981) discussed formative evaluation at the British Museum (Natural History). However, other than Shettel (1973) and Screven (1974), few put all the parts together and saw them as a systems approach to museum exhibits and educational programs.

Miles (1982) synthesized all the various parts after looking for ten years. In creating a large exhibit for an entire exhibit hall at the British Museum (Natural History), Miles and his colleagues applied the various parts of instructional design in a systems approach. They published their experiences and advice in *The Design of Educational Exhibits*. In the preface, Miles concluded that "the failures of the past had not been so much failures of people, but failures in the information available to them." In 1984 the American Association of Museums (1984)published

Museums for a New Century: A Report of the Commission of Museums for a New Century. This work was a major examination of the museum field in the United States. It recommended that museums specifically improve the educational aspects of their exhibits and educational programs. It mentioned contributions of Screven and the application of instructional design. Further, it recommended that museums interact more with their constituencies.

From the date that Museums for a New Century (1984) was published, museums have expressed much commitment to facilitating learning and communicating with their constituencies. However, there are still a number of "failures" due, as Miles (1982) put it, to failures of information and not people. To cite a simple and recent example: a major museum's reinstallation of a permanent exhibit included labels at ankle level which were less than one fourth inch high and which were difficult to read even on one's knees. Next to the labels was a sign which apologized for the problem, which was "being evaluated". Such occurrences raise a variety of concerns when museums talk of using computers and interactive video in their exhibits and educational programs.

The traditional educational program and exhibits at museums were created by experts for experts. There was little attempt to communicate on the level of their constituency. There was frequent use of the outstanding or very rare example rather than the "model case" artifact which would teach much more about the concept. Labels were written in a meaningless fashion, for example, "chair, wood, stained" which are all obvious to all who viewed the chair. There was little attempt to explain why "the chair" had meaning.

Major Findings, Conclusions, and Recommendations

Major Findings

The objective of the "Exhibit Development Survey 1986" was to produce a description of the exhibit development process -- the methods museums use to develop their exhibits -- at major American museums. It was hoped that this would also reveal if these museums are using the instructional development process to develop exhibits, and if so, to what extent. Though the exhibit development process is an exceedingly complex one, the results of the survey yield a fairly concise picture of the means by which museums go about the process of developing exhibits.

The survey results indicated that the development of exhibits has become the primary area of concern for major American museums. The museums view education as their second most important area of involvement. Further, they view their exhibits as their most important educational activity. Thus, the exhibit development process, in which they concentrate much of their energies and resources, is their main avenue to educate to the general public.

This configuration of priorities produces a symbiotic relationship between exhibits and education: The museums' first priority is their most important means of attaining their second priority. These priorities must be questioned, however, when one examines other survey responses. For example, 24% of respondents said their institution does not consider the needs or preferences of their audience when developing exhibits. Seventy-four percent of responding institutions do not use a data bank of their audiences' characteristics when selecting exhibit topics. Fifty percent of responding institutions do not try to assess what their audiences learn from their exhibits. These figures would seem to indicate a lack of emphasis on the educational value of exhibits by a number of major museums.

However, it should be noted that many of the major museums extol to the importance of educationally effective exhibits. Indeed, the survey responses taken as a whole, plus respondents' additional comments, indicate that most major museums consider their exhibits to be a serious educational endeavor. Further, the exhibit development process is an area which has undergone many recent changes -- an area museum professionals take seriously and are eager to improve.

The survey also revealed that there has been a major shift towards using instructional development as the basis of the exhibit development process. Fully one third of major American museums indicated they are using instructional development to create exhibits. Thirty-nine percent of respondents indicated their institution included an instructional developer on their exhibit development team, and another 32% felt their institution should utilize instructional developers for this purpose. Still other institutions are using individual instructional development techniques as part of their exhibit development process, but are not taking advantage of the benefits of the overall systematic process. It is unclear whether these institutions are aware of such benefits, or whether they do not know how to implement the instructional development process.

These statistics are not just remarkable because of the percentage of responding institutions utilizing instructional development, or advocating the use of instructional development, or advocating the use of instructional developers, but also because this development has largely been overlooked in the museum literature. Further, considering the primary role of exhibits in the museum profession, it is striking that there have been few studies of the overall exhibit development process, and in particular, the recent utilization of instructional development in this process.

Another implication, which the results of the survey raised, concerns the separation of the education and exhibit divisions in some museums. While specific survey questions did not address this point, a number of respondents' comments indicated this was the case at their institution. In several cases, their comments indicated that exhibits were produced by a "curatorial division", or the "exhibits division", and then the responsibility for using the exhibit in an educational manner became the responsibility of the education division. Further, a number of these respondents appeared not to know or care, if anyone learned anything from their exhibit. They suggested this was the responsibility of the education division -- or at least the education division staff should be the ones to ascertain this. However, it would seem to follow that if those persons developing an exhibit are not vitally interested in its educational effectiveness, then the importance of this aspect of the exhibit is diminished for those persons.

Of necessity, the scope of this survey was limited by the expansive nature of the exhibit development process. Each of the various procedures of the exhibit development process needs to be examined in greater depth, but the overall process must be kept in view at the same time. These areas of concern -- from selecting an exhibit topic to evaluation of the educational effectiveness of an exhibit -- are not isolated procedures; they are all part of the exhibit development process and should be examined in that light.

Conclusions

The conclusions of the survey are reported in this section. For this present discussion, the content of the survey has been divided into ten main areas of concern, plus an additional section addressing respondent remarks to several open-ended questions.

Demographics of Respondents and Responding Institutions.

The response to the survey was quite adequate with approximately 24% of all museums in the population of major American museums responding to the survey. Further, the responding institutions represent a broad cross-section of major museums with 32% being art museums, 18% history museums, 32% science museums, 5% natural history/ anthropology museums, 8% children's museums, and 5% general museums. The respondents, who actually filled-out the survey for their particular institution, fell into five categories: administrative for the institution -- 34%, curators -- 8%, educators -- 13%, administrative for exhibits division -- 8%, and staff members from the exhibit division -- 8%. Ninety-five percent of the responding institutions said they served an audience of the general population.

Institutional Priorities.

The responding museums ranked exhibits as their most important priority with education a close second. These two activities were ranked, by far, as these museums' most important priorities. Exhibits, among various museum activities, were also considered to be these institutions' most important form of educational activity. However, they believed that a somewhat greater number of their visitors came to their institutions for "recreational reasons," rather than for educational ones. This is further accentuated by the fact that the respondents also ranked "social reasons" as an additional important reason why visitors came to their museums.

Development of Exhibits.

The responses concerning the steps used to develop an exhibit indicate that the process is a lengthy one, utilizing a wide variety of procedures to complete the exhibit development process. Further, the responses seem to indicate that the length of time devoted to the development process would allow enough time for adopting innovative strategies, or the use of instructional development procedures if further study indicated that such methods are effective. While some would argue that such innovations would lengthen the exhibit development process, it could also be argued that the use of a systematic development process could make the process more efficient.

The survey also revealed that the steps museums use to develop exhibits lean toward what might be termed "front-end" analysis -- those activities which get the exhibit started. However, their utilization rate was considerably lower for those aspects of exhibit development which would reveal how successful their work had been, or would reveal ways to improve their exhibits.

From the comments of those respondents who do utilize instructional development, it was clearly emphasized that formative evaluation is believed to be a very important/crucial step in the exhibit development process which produces improvements in their exhibits. Trial runs of exhibits were utilized by a few institutions, but this step seems somewhat neglected by even those who indicate that they utilize instructional development. Perhaps through formative evaluation, individual segments of exhibits are put through what might be termed partial "trial runs."

It should also be noted that it may have been better if the survey had used the term "trial runs of exhibit mockups", or had used this phrase as an additional category/response choice. Completing "trial runs", or formative evaluation, of exhibit mockups would be more likely to occur than trial runs of finished exhibits. This would also be a more realistic and practical step before actually conducting a trial run of a full-scale, installed exhibit.

Consideration Given Visitors/Audience.

Survey responses concerning the ways that responding institutions consider their visitors/audience in the exhibit development process indicate that a variety of techniques are used to address this vital concern. Seventy-six percent of the respondents felt their institution's exhibits were responsive to its audience's needs or preferences. However, the overall picture is not quite so positive, considering 24% of respondents do not consider their audience's needs or preferences at all when developing exhibits.

Additionally, respondents indicated that assessing the learning style of their audience and utilizing this information in the exhibit development process was the method of operation in about half of the responding institutions. Further, an even larger number of these institutions designed their exhibits so that their visitors/audience are guided through an exhibit to encounter various sections in the order in which various "pieces" of information are to be learned.

Use of Instructional Objectives to Develop Exhibits.

A rather high percentage of institutions revealed that they were using instructional objectives to develop their exhibits. However, the majority of respondents who used instructional objectives also indicated that they were stated in general terms such as "appreciate" or "understand". Only 8% indicated their instructional objectives were stated in behavioral terms -- or what the audience "will be able to do" after visiting an exhibit. (It should be noted that this low figure for the use of behavioral objectives is in conflict with the number of institutions [34%] claiming to use instructional development to develop exhibits. This might possibly indicate that they are still in a transitional stage to the full use of the instructional development process of exhibit development.)

Those institutions that do use instructional objectives indicated that the objectives were developed during the entire exhibit development process. Their responses indicated that many of these institutions begin to develop the instructional objectives for their exhibits early in the exhibit development process. However, these figures also seem to indicate that the respondents are modifying their objectives as their exhibits take on more concrete dimensions or parameters. At first glance the 11% who develop their instructional objectives after an exhibit is finished would seem not to be making proper use of such objectives. However, there may be legitimate cases where additional instructional outcomes are identified after an exhibit is completed. It would then be appropriate to develop instructional objectives and integral or peripheral educational aspects for an exhibit to accomplish these objectives. It would take an exhaustive study of a large number of individual museums and how they actually utilize instructional objectives to be able to generalize about the true value of instructional objectives in the exhibit development process and when they would be best written.

Integral and Peripheral Educational Aspects of Exhibits.

In the category of integral/peripheral educational exhibit aspects, the designer of the survey expected to receive mainly approximate answers to the survey questions, assuming that museums would not keep statistics on funding and staff time expended on these aspects of an exhibit. The respondents indicated that these questions were difficult to answer and that they frequently answered them approximately. Their answers indicated that both integral and peripheral educational aspects of an exhibit are considered early in the exhibit development process -- though they begin to consider peripheral aspects at a later date than the integral aspects. The responses would appear to indicate that these activities respond to the ebb and flow of the exhibit development process. That is, if the exhibit begins to take some new direction, or as the direction of the exhibit becomes more clear, corresponding integral and peripheral educational aspects of the exhibit are developed. However, without further study, it is impossible to tell whether those educational aspects (both integral and peripheral) which are considered late in the exhibit development process, are just "tacked on" to the exhibit or if they serve a truly educational function of the exhibit . The respondents indicated that considerably more of their educational efforts, staff time, and budget were devoted to the integral educational aspects of their exhibits, than were devoted to the peripheral educational aspects.

Evaluation Procedures for Exhibit Development.

The evaluation procedures used by the responding institutions to investigate the effectiveness of their exhibits include a wide range of techniques, from testing and trial runs of exhibits to the use of interviews, questionnaires and outside consultants. However, three-fourths of responding institutions did not use trial runs of their exhibits before the exhibits were completed and open to the public. Those institutions which did not use trial runs indicated a lack of time, funds, or staff to do so. Others had not considered this option, or felt it was unnecessary.

Concerning learning produced by an exhibit, half of the responding institutions do not try to assess what, if anything, their audience learns from their exhibits. This statistic stands out starkly against the high priority given education -- education through exhibits -- by these same institutions. While the learning taking place may not be "zero", these institutions would be hard pressed to demonstrate to many educators that this is not the case. However, they do utilize a number of other strategies which would help to indicate what their audience's reaction is to their exhibits.

When asked to rank what indicated a successful exhibit, the respondents indicated that the most important indication of a successful exhibit was large crowds visiting the exhibit. The response of the critics was another important indication of the success of an exhibit. Questionnaires were also frequently used as a way of learning in an exhibit was successful. Interviews were used even less frequently, and testing ranked very low as a means of learning if an exhibit had been successful.

Forty-five percent of the survey sample did not respond to this question about exhibit evaluation. One is tempted to draw the conclusion that those who did not respond believe that their exhibits are not successful or they do not employ a strategy which would let them know if their exhibits are successful.

Respondents indicated that nearly two-thirds of the institutions actively seek feedback or critical reaction from their audience -- after the exhibit is completed. Of those institutions that do seek feedback or critical response, the majority indicated that such feedback would affect their current exhibits. An even larger percentage of respondents indicated that this feedback would affect future exhibits.

Coordinating the Exhibit Development Process.

Concerning the question of coordinating the exhibit development process to insure its educational soundness, 63% of the respondents indicated that their institution had an official staff position responsible for this duty. Those respondents who had such a position further indicated these positions ranged from administrative for the museum to administrative for the exhibits division. The position was also held by curators, staff from education/public affairs, and exhibit designers. Several museums indicated they used a team approach for this function.

For those institutions which did not have such an official position, the responsibility for coordinating the work on an exhibit usually was assumed by one of four main museum staff positions: a curator, an educator, the museum director, or an exhibit designer.

Current Exhibit Development Process.

The responses to the survey point to a flurry of activity and interest in changing and improving the exhibit development process. In addition, the survey results indicate that the exhibit development process at most of the respondents' institutions has been adopted fairly recently. In the last 5 years, 45% of the responding institutions have adopted a new process to develop their exhibits. Twenty-four percent have changed their process in the last 6 to 10 years and 13% have changed in the last 11 to 15 years. Further, it should be noted that 34% of responding museums utilize instructional development for their exhibit development process. Additionally, 6% of respondents said they were using some instructional development procedures or similar procedures to develop exhibits. The comments of respondents indicated in several instances that they had recently changed to the use of instructional development, were in the process of doing so, or were looking into the possibility of such a change. Those institutions which were using instructional development indicated that 18% felt this process of developing exhibits was very successful, 13% felt it was successful, and 3% described it as "common sense".

Fifty-six percent of respondents indicated that their institutions were not using instructional development for their exhibit development process. Of those institutions that were not utilizing instructional development, 24% cited "other" reasons for not doing so. In general, their comments implied that their exhibit development process was moving in that direction (use of instructional development), or that time, limited resources or other limitations prevented the use of instructional development. It also should be noted that 21% of respondents were "unfamiliar with this (instructional development)process".

Concerning whether the exhibit development process at their institution should be restructured, 29% of respondents felt the exhibit development process was fine at their institution, while 18% felt it needed to be restructured. The majority of the respondents answered "other" concerning this question. Their comments under the "other" section indicated that they feel the exhibit development process is in a state of flux, of "constant growth and reworking" -- which would seem to indicate that many of those who chose "other" are interested in ways to improve the the exhibit development process.

Instructional Developers in the Exhibit Development Process.

Respondents indicated that at 74% of their institutions some members of their staff were familiar with the procedures of instructional development. Additionally, 45% of respondents indicated there were instructional developers on the staff of their institution. Respondents whose institutions had instructional developers on their staff also indicated that these instructional developers worked in a variety of positions in the museum; from curatorial departments, to exhibits, to education/interpretation, etc. Their duties were also quite diverse, including coordinating exhibits, acting as evaluation/audience experts, planning exhibits, planning programs, working with curators, and acting as exhibit team members.

Regarding the role of instructional developers in the exhibit development process, it should be noted that 39% of the responding institutions have an instructional developer as a member of their exhibit development team. An additional 32% of the respondents believe their institution should include an instructional developer as a member of their exhibit development team. This would seem to indicate that 71% of responding institutions believe that instructional developers should be a part of the exhibit development team. Further, when considering this 71% figure, it should be kept in mind that 21% of respondents were not familiar with the instructional development process.

While a large percentage (71%) of respondents indicated that their institutions currently had or should have an instructional developer as a member of their exhibit development team, 55% felt that instructional developers should not coordinate the work of those developing an exhibit. Those that felt an instructional developer should not coordinate the exhibit development process indicated that this was more appropriately done by either an administrator, a curator, an exhibit team, an exhibit committee, an exhibit coordinator, a project director, or by assigning this responsibility to various "owners". However, 11% of respondents indicated that at their institutions instructional developers were coordinating their exhibit development process, and another 18% felt this should be the case.

Additional Respondent Comments.

Several open-ended questions at the end of the survey allowed the respondents to discuss areas which the survey had not touched upon. One respondent wondered if exhibition design was based on educational or aesthetic considerations; what percentage of exhibits were organized in-house and what percentage were traveling exhibits; and they wanted to know if there was anyone who was charged with the responsibility of insuring a coordinated and fiscally responsible project. Another respondent felt the question of what exhibit components were specifically budgeted should have been investigated, stating that, at his institution, education/ peripherals were not specifically budgeted. Another of his questions considered whether museums are generally happy with their exhibit development process; he answered "no" for his institution. And finally, he wanted to know the most serious problem with museums' current exhibit development processes. He indicated that at his institution it was: "too time-consuming and no one person in charge".

Another respondent wondered if museums were developing long-term plans -- such as a 5 year plan of exhibitions. Their institution was then developing a 3 to 5 year plan. They also wondered if other institutions had developed written statements of purpose for exhibits. Again, their institution was developing such a statement. One last respondent wanted to ask, "what percent of the dynamics of exhibit design and fabrication are conducted along formal lines, with guidelines; and what percent is from successful, informal routine"? Their answer indicated that they felt a museum should be "flexible enough (and professional enough) to apply the techniques, personnel and consultants as needed" to respond to the different dynamics of each exhibit.

Recommendations

It is not the purpose of this study to examine the "correctness" of the priorities of major American museums. However, it is clear that these institutions have identified exhibits, education, and education through exhibits as their main priorities. If it is thus safe to assume that they are interested in developing educational exhibits, then the survey results would seem to indicate that there is considerably more which many museums can do toward that end -- in particular, in examining and improving the exhibit development process.

As museums develop exhibits, they could make much more extensive use of a number of strategies to make their exhibits more educationally effective. Certainly more could be done to consider the museum visitor in the selection of exhibit topics, both regarding what topics serve visitor needs and preferences and what topics visitors have the requisite knowledge to understand.

In the actual exhibit development process, such activities as needs assessment, goal setting, and writing of instructional objectives can all be utilized to greater degree. This is particularly true of instructional objectives, which should be used much more frequently and also should be more specific and stated in behavioral terms. Further, the instructional objectives should be linked to various evaluation processes to insure that the objectives are being reached.

Efforts to evaluate the effectiveness of exhibits need to be greatly increased. Formative evaluation needs to be made the norm in the exhibit development process, while much greater use of trial runs of exhibit mock-ups and trial runs of completed exhibits needs to be madeto aid in the development of effective exhibits. The use of trial runs may initially add some time to the exhibit development process, but with experience this time difference should be shortened and also compensated by the increase in the quality of the exhibits. Indeed, it seems plausible that museums would learn a "generalizable process" of using trial runs for formative evaluation of exhibits, or at least for specific types of exhibits, which would speed up this process and make it very efficient.

It is advisable that those museums which do not try to assess the educational effectiveness of their exhibits begin to do so as normal museum practice. For those museums which do assess learning, efforts need to be intensified to increase learning effectiveness through improved systematic and effective exhibit development processes -- including assessments of learner/visitor outcomes. For all museums, the successful exhibit must come to be seen as an educationally effective one, not just one attended by large crowds or praised by critics. Hopefully all three criteria can be fulfilled.

Museums need to make much greater use of feedback mechanisms and use the feedback they receive as a means of improving future exhibits. They must learn from both their successes and failures and incorporate what they have learned into the process by which they develop exhibits. Museums must become more aggressive in seeking feedback. They need to know not only whether visitors "liked" an exhibit, but also whether numerous other factors were well executed-- such as the effectiveness of lighting, use of color, shape and size of the exhibit space, use of labels or other textual materials -- that would have a significant impact on the effectiveness of an exhibit. And they need to know if anyone learned from the exhibit.

How the exhibit development process is coordinated or who performs this vital function remains open to question. The responding institutions appeared to vary greatly in this aspect of the exhibit development process, and it appears that these museums follow a variety of strategies. Though the title of the coordinator may vary, it appears important that there be one person or a small group of persons which controls the process. If exhibits are to be educational, then it would seem to follow that the person who coordinates the exhibit development process should also view this as his/her purpose -- to bring together the frequently disparate members of the team to create educationally effective exhibits.

Both the extent of the use of instructional development for exhibit development (34%), and the number of museums using instructional developers in their exhibit development process or respondents who believe their institutions should use instructional developers in their exhibit development process (71%) would seem to indicate that major American museums should begin to examine the usefulness of this technology in an organized manner. If additional museums are to adopt instructional development, it should be done in full knowledge of what this entails, and it should be done in a manner that would be most likely to insure its greatest success. This can only be done through further study of the exhibit development process and the use of instructional development.

This is no small task as very little research has been conducted on the overall exhibit development process. And while it may appear to a particular museum's staff that it is developing excellent exhibits, many museums appear to have no way to determine this or compare their exhibit development process to another one. Thus the establishment of a central organization charged with the investigation of the exhibit development process would seem appropriate -- to investigate such issues as the adoption of instructional development and its adaptation into a museum technology.

The major museums in America need to establish a systematic means to allow research into their activities. Currently there is very little centralized information which would allow researchers to investigate the exhibit development process or other issues of importance to museums. The American Association of Museum's Official Museum Directory, while serving a useful purpose, does not contain the type or quantity of information which researchers would need. The Institute of Museum Services, which provided the statistics (museums by discipline and budget size) which made this current research possible, does not have all the information needed nor do they have the time or resources to provide such information to more than just the occasional researcher.

Further, an experimental exhibit "center" could be established to investigate various ways of developing and implementing successful exhibits -- information in which major museums can share. This might be done by each of the various museum disciplines, as a consortium; and it could be done by producing experimental exhibits at a number of museums on a rotating basis, with member museums sharing in the knowledge acquired and the extra costs of such experimental exhibits. In this way instructional development or various forms of hardware, such as interactive video, can be tried out on a cost sharing basis in innovative ways. In this manner major museums can begin to share in the advances in educational research and technology without each museum individually bearing the costs and failures such change and experimentation would entail.

In addition, it would not be just the major museums which would benefit from such a move. Certainly their visitors would benefit, but also the thousands of smaller museums and historical societies might also receive great benefits. Not only are the nation's major museums role models for these smaller institutions, but they also provide them with expertise and training for a variety of museum activities. In particular, exhibit development would seem to be a primary area where the major museums might give assistance and guidance to these small institutions. Perhaps some of the major museums, which are currently utilizing the instructional development process successfully for exhibit development, should offer consultation services to other museums which would like to begin using instructional development.

Museums also need to examine their hiring practices. Those museums which require advanced degrees in a subject field -- such as art history for an art museum, or an advanced science degree for a science museum -- for museum education positions seem to be overlooking the real role of the museum educator. Museums, in their zealousness to hire educators who know their subject, overlook whether these persons can design instruction. This is not to say that museums should not seek educators knowledgeable about the museum's discipline, but that their ability to develop effective instruction -- be it lecture, exhibit, printed matter, or outreach programs -- must be of, at least, equal importance. The competent instructional developer, with some subject expertise, can develop effective instruction in cooperation with other subject experts on the museum staff. The person with a particular subject matter background can perhaps provide the subject content, but this knowledge does not indicate that they have the skills to design "instructional systems" by adapting this content to form the basis of an exhibit or other educational program.

A task force should be created to foster the development of a museum technology. While instructional development (educational technology) might well serve as the basis for such a technology, this should be verified by additional research and field trials at major museums. A systematic and verifiable process to unify and facilitate museum functions is needed.

Indeed, an excellent start to create a museum technology was made by a number of museum professionals at the British Museum (Natural History) for use in developing exhibits for their "new" Hall of Human Biology. They realized that if they were to create educationally effective exhibits, they would need to adopt a different exhibit development process. Thus, they systematically set out to develop a museum technology for the development of exhibits. The accumulated knowledge and experience from their pioneering work is available in the book, *The Design of Educational Exhibits*, edited by R. S. Miles. It is by far the most comprehensive work on the development of exhibits, one which incorporates instructional development (educational technology) and lays the foundation for a museum technology. This work should be studied by all museum professionals, particularly those involved in the development of exhibits.

Additionally, instructional development may be the unifying technology for which museum professionals have been looking to make the museum profession more cohesive and identifiable as a profession. It would allow, in combination with other concerns addressed in university museum studies programs, a fuller communication across the various job titles in the profession.

Regarding the usefulness and value of instructional development to the museum profession, it would seem appropriate to recall the remarks of three survey respondents. One respondent, whose institution did not use instructional development, remarked that, "the survey showed naive understanding of how museums and other organizations operate." In fact, just the opposite would seem to be the case. This particular respondent does not grasp that the issues reflected in the survey are the concern of many organizations and American corporations (and 34% of major American museums) as more and more of them adopt the use of instructional development. Major corporations, which expend millions of dollars per year on education/staff development/training can not afford to do this without verifiable outcomes from these educational/training expenditures, as they can not for expenditures in other areas.

Why should museums be any different? The answer is, they should not. Just because museum exhibits offer non-directed learning is no excuse for a museum to abrogate its mandate to educate its audience; that is, provide an experience/exhibit which produces learning which is verifiable. This means, if a visitor to an exhibit actively participates in the exhibit, a certain amount of learning should take place. For this learning to be verifiable, the exhibit must have objectives for what is to be learned from the exhibit. If these objectives codify what is to be learned from an exhibit then they are instructional objectives. And if these instructional objectives are to be verified they must be stated in behavioral terms, because learning can not be measured/verified unless a change in behavior can be demonstrated.

Another respondent, from a museum that utilizes instructional development, indicated one"problem" with its use. She pointed out that when a new person joined their staff, they had to be trained in the use of the instructional development process. At first glance, this sounds like a drawback to its use, at least until many more museum professionals are trained in this discipline. However, museums utilizing instructional development should ease and facilitate the transition of current employees and new employees to the use of instructional development with orientation training, as corporations train and orientate their new employees to their system of operation. Such training for all employees would facilitate communication between museum professionals at different museums and between disciplines -- educators, curators, designers, administrators, researchers and other staff members -- within the same museum. This is not to say that curators or exhibit designers need a "degree" in instructional development, but instead that they would need an overview/in-service training course to be knowledgeable about instructional development to begin to work effectively in such a system. Eventually, if broadly adopted, a systematic, professional process -- like instructional development -- would ease the orientation of new museum workers, or a museum worker who accepts a position at another museum.

A third survey respondent indicated that they felt the survey was comprised of "educational jargon and philosophy". While considerable educational "jargon" -- learning style, instructional objective, the word "education", integral and peripheral educational aspects, "learns," instructional development/developer -- was mentioned, much of the survey was also comprised of "jargon" which describes a systematic means of attaining objectives. The officers of many corporations would not think such terms as needs assessment, summative evaluation, formative evaluation, audience served, ranking priorities, development time, trial runs, utilization of feedback, surveys, questionnaires, or interviews to be educational jargon. They would recognize them as elements of a systematic process which guides an organization. Should museum professionals be any less familiar with the means of the systematic operation of an organization? Or, for that matter, should they be any less familiar with a variety of educational jargon since the stated mission of most museums is education? This respondent's comment points to a lack of understanding of the benefits of working in a systematic manner and the lingering denial of some museum professionals of museums' educational role and orientation.

Museums which are currently using instructional development for exhibit development, as well as those planning such an introduction, should develop comprehensive in-service training to orient current employees and future employees to the use of instructional development in a museum setting. Such a step is crucial to the successful utilization or adoption of instructional development. In some cases, it may be necessary for certain individuals to assume the role of "change agent" within their particular institution. However, this role is not without risk, frustration, or aggravation. Nor is everyone in a position to attempt this beyond their own immediate sphere of influence.

The responses of those museums which are utilizing the instructional development process indicate that many of them are utilizing a number of individual instructional development procedures. It is likely that these museums would find the adoption of this technology a fairly easy task to accomplish since they are already accustomed to various aspects of the instructional development process. Further, their use of some procedures of the instructional development process can only be enhanced by the full spectrum of the instructional development process. These individual procedures would then be more successful and more integrated into the overall operation of the museum.

The response to the survey would seem to indicate the museum profession's interest and concern about the process of developing exhibits. They appear to want to improve this process. However, the museum profession lacks any organized, systematic effort to examine exhibit development. This study has revealed a lack of research in this area and the unavailability of statistical information on American museums. Both of these situations should be remedied by a museum studies clearing house supported by government funding and a consortium of major museums. This will benefit the major museums, as well as the hundreds of smaller museums which model their practices after these leading institutions.

It is of particular interest that 39% of major American museums are currently using instructional developers in the exhibit development process, and that another 32% of the respondents believe their institution should be using instructional developers to develop exhibits. These statistics are puzzling in view of the infrequent number of times instructional development is mentioned in the museum literature. Further, the fact that 34% of museums are using instructional development to develop exhibits indicates that the role of this process in exhibit development should be examined by the museum profession and at least considered in university museum studies programs. If, as the survey respondents indicate, it is a successful way of developing exhibits, then further measures should be taken to examine and possibly adopt and adapt instructional development as a museum technology.

References

American Association of Museums. (1969). America's museums: The Belmont Report. Washington, D.C.: Author.

American Association of Museums. (1984). *Museums for a new century: A report of the Commission on Museums for a New Century*. Washington, D.C.: Author.

Cameron, D. F. (1968). A viewpoint: The museum as a communication system and implications for museum education. *Curator, 11*(1), 33-40.

Griggs, S. (1981). Formative evaluation of exhibits at the British Museum (Natural History). *Curator*, 24(3), 189-201.

Miles, R. S., Alt, M. B., Gosling, D. c., Lewis, B. N., Tout, A. F. (1982). The design of educational exhibits. London: George Allen & Unwin.

Screven, C. G. (1974, January/February). Learning and exhibits: Instructional design. *Museum News*, pp.67-75.

Shettel, H. H. (1973, September). Exhibits: Art form or educational medium? *Museum News*, pp. 32-41.

Title:

Distance Education in Taiwan: A Model Validated

Authors:

Mei-Yau Shih
Susan M. Zvacek

Research and Theory in Distance Education

Distance education systems throughout the world have begun to exploit the technological advances that are shrinking our planet and making geographical separation all but irrelevant. As distance education continues to grow, the need for credible and rigorous research on designing, implementing, managing, and evaluating these programs increases. Scholarly research, in turn, rests upon our understanding and application of theory.

If, as Holmberg (1985) suggested, a theory is "a set of hypotheses logically related to one another in explaining and predicting occurrences" (p. 2), then Wedemeyer's (1973) assertion and Keegan's (1990) concurrence that distance education has failed "to develop a theory related to the mainstream of educational thought and practice" (p. 52) holds serious implications for researchers in this field. Baath (1982) argued that distance education would proceed from the "deliberate application of relevant educational theories or models" (p. 37). How, then, to determine what theories are "relevant" to distance education and, once determined, how should those theories and models guide the "explaining and predicting" process?

The importance of sound, theoretically-based research cannot be overemphasized. This emphasis on theory is what will enable researchers to "reassess, if necessary in the most painful of manners, those funny elements of the fast-developing lore of distance education with the object of determining which are really crucial to the success of our mission and which are simply there because we happen to like them" (Griew, 1982, p. 191). A model designed for examining distance education realistically -- as opposed to idealistically -- is described in the following paragraphs.

The Triad Perspective Model of Distance Education

The Triad Perspective Model of Distance Education (TPMDE) was developed as part of a research effort attempting to describe, in an organized and systematic way, distance education activities in Zimbabwe (Zvacek, 1990). This model guides the researcher in developing research questions, gathering data, and producing a comprehensive description of the distance education program being examined.

The "triad" label refers to the three theoretical perspectives around which the model was developed. The first is curriculum development theory and is based on Tyler's "four questions" for planning instructional programs (Tyler, 1949). The second perspective is systems theory, principally the theoretical models developed by Banathy (1973). The third element of the triad is based on Rogers' work on adoption and diffusion of innovations (1983). These three theoretical perspectives were integrated and organized graphically to create four "levels" or planes of reference from which the researcher could study a distance education program: environment, structure, process, and evaluation. (For a complete description of the TPMDE and its development, see Zvacek, 1990.)

In an effort to validate and improve the TPMDE, it was used to guide research efforts in a project to study the Taiwan National Open University

(Shih, 1990). The graphic representation of the TPMDE was modified in order to emphasize the dynamic nature of the model and highlight its "three-dimensional" nature (the original representation of the model was built using Tinker Toys). (See Figure 1.) This new representation, the "Glassview TPMDE" also helped to define the relationships among the theoretical strands of the model.

The primary purpose of the Triad Perspective Model is to guide researchers in creating comprehensive descriptions of distance education systems. This descriptive feature could act as a first step in evaluating a program or tracking progress over time. Another option is to use only one of the four levels of the model in examining a particular program -- environment, for example, to better understand how the distance education program influences and is influenced by its cultural or political setting.

Another possible function of the TPMDE is as a design or planning tool. The perspectives used in studying a program to better understand it lend themselves to planning a system strengthened by theoretical underpinnings. The research questions would be design considerations related to the four reference planes of environment, structure, process, and evaluation. These potential applications, and others, will be examined in future research studies.

Distance Education in Taiwan

Distance education presents a new challenge for Taiwan. This perception of education is different from traditional education in Taiwan in two ways: first, in the concept of equal opportunity for education and, second, in the process of instructional development.

With limited natural resources in Taiwan, economic and political success is perceived to be rooted in equal opportunity for education. A higher education background is crucial for a job in a government office or in private business and is still considered by many a measure of social status in Taiwan (Wilson, 1970). The need for more higher education opportunities has been acknowledged in recent years and subsequent projects have been successful in establishing more universities and colleges. In 1986, Taiwan established a National Open University which borrowed the distance learning/teaching concepts from the British Open University and the structure of distance education from Japan; it has attempted to narrow the gap between desired outcome--more opportunity for higher education--and the actual situation.

The application of distance education in Taiwan has brought with it the new process of instructional development. The methods of instructional delivery in distance education involve the application of various media with teachers and students in different locations; teaching occurs at a variety of sites simultaneously or at different times. This presentation format is difficult to accept in the traditional Taiwanese educational environment because using TV, video, computer, or phone lines to deliver a lesson is novel to most teachers and students. Chinese teachers and students are raised to respect strict discipline in the classroom and to compete with other students in overt performance, which includes interaction with teachers and peers, and appropriate behavior in the classroom. The classroom atmosphere is set in kindergarten and maintained through college. Therefore, primary criticisms of distance education arise from a perceived lack of interaction and surveillance in the instructional setting and from what is seen to be a more loosely structured teaching-learning environment.

The National Open University in Taiwan was established in 1986 as a "different type" of university (Cheng, 1990). The first proposal for the establishment of NOU was sent to Legislature Yuan in 1983 and a distance education committee went abroad to Britain, South Africa, and Japan to observe their distance education systems and applications. The goals and functions of NOU have been repeatedly debated by legislators and educators in Taiwan in past years. The NOU remains a cooperative effort of educators, administrators, and government agents in adopting a new techology within an old society.

TPMDE in Taiwan

By adopting the TPMDE's four levels of reference in studying a new distance education program such as NOU in Taiwan, the program itself was able to be presented within a social context, as well as a more objective viewpoint. Since the nature of TPMDE is a perspective model, it could present a general understanding through a telescoping approach, that is, the whole picture of a program, instead of trivia, would be shown. In that context, distance education at the NOU, Taiwan, was examined from social/environmental perspectives, following the framework. First, the objectives and goals of NOU were identified and how the NOU approached its goals was also revealed. Expectations and perceptions of NOU by other society members and students were also studied and revealed differences between NOU students and conventional university students. NOU students' self-perception scores were higher than the students' self-perception from regular universities. This may suggest that the NOU students perceived themselves as more valued because they were able to retake the college courses after they had missed the first chance earlier and they may feel more self-determined in their choice of going back to the university. Further study should concentrate on the examination of attributes which caused this difference. According to this study, the general public's perception of NOU was vague as to NOU' s role in Taiwan society and higher education. Even the high percentage of citizens who recognized NOU revealed perceptions which suggested there still exists a gap between what NOU is and what the public thinks it is.

The second level in the TPMDE then allowed one to investigate more closely the program structure and innovation itself--the components that functioned in the NOU and moved the NOU toward its goals in the past three and one-half years. In studying the interaction of these components within NOU, suggestions then can be made to add, drop, or modify components. Components such as teacher training and public relations were suggested additions to the NOU structure and the counseling division, which has not operated effectively, was seen as one division best abolished or modified.

The third level of the model, process, is concerned with the program's development over time. The NOU has evolved over the past three and one-half years from 11 sites to 12 sites and student population had risen to 55,653 by Spring 1990. The number of college level courses offered increased to 44 by Spring 1990.

Evaluation was the fourth level of the model. A regular evaluation procedure did not exist in NOU, however, there were administrative levels of communication and faculty communication levels which were based on academic concerns. The highest marks for NOU given by students concerned the opportunity for adult training, flexibility of course selections, and expansion of the knowledge horizon. Most faculty members also appreciated the challenge of distance teaching, however, they felt the lack of a support system in preparation of instruction did impede the effectiveness of the overall program.

Summary

Several recommendations for further research are suggested for the NOU. First, effective media utilization needs to be studied. The study should include the most suitable media for a given content area and type of audiences. Careful measurement of TV effectiveness and cost-effectiveness in various content areas would also help to determine appropriate usage of TV channels to deliver the instruction.

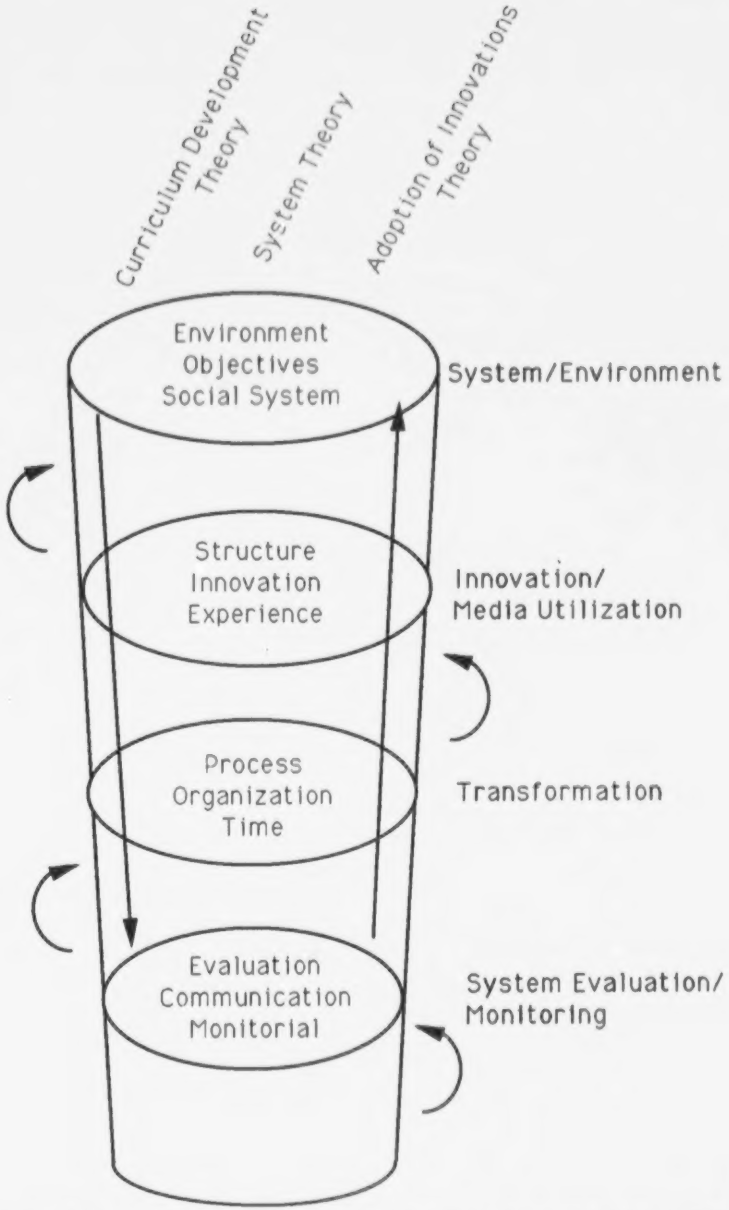
Second, a study of motivation of NOU students and the correlation of the motivation and students' performance, as well as the correlations between face-to-face instructional meeting attendance and students' performance should be implemented. This study would help locate the students' foremost motivation in learning and the influences of face-to-face instruction in distant learning. By doing so, revision of overall coursework design and development in NOU may be made.

Third, a regression study on factors which correlate with job satisfaction in NOU staff and faculty would be helpful in understanding the program. This study could be followed by a comparison study of job satisfaction between NOU staff, faculty, and conventional university staff and faculty.

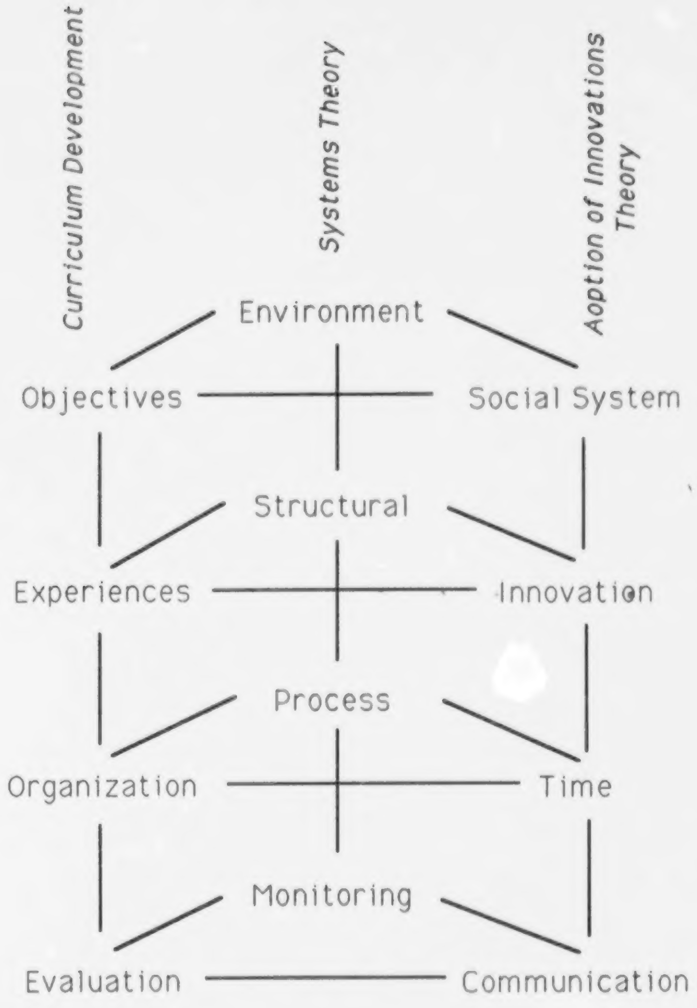
Further research utilizing the TPMDE is also recommended. By providing a comprehensive and theoretical view of the NOU in Taiwan, this study served to bridge the gap between a theoretical model and application of a real life program. The results drawn from the findings concluded that distance education is a complex team effort that requires careful planning based on set goals and objectives. Cooperation among components of the distance program is also critical to its effectiveness. Most importantly, the distance system has its own unique social context and may require some adjustments to fit into each individual environment. The adjustments could include delivery format, instructional presentation, instructional design, implementation, and learner characteristics.

Distance education is seen as a potential "promised land" in Taiwan which will provide equal access and economical higher education to more people in various levels of education and trainings. It is a means to an end to facilitate learning and human performance, not withstanding the supporting theories critical of distance education's complex operation. The primary purpose of this study was to link the social phenomenona, instructional design efforts, and application into a theoretical explanation of distance education operations in Taiwan.

Globally, distance education will profit from a strong research base undergirded by theoretical knowledge. Only then will researchers be able to "explain and predict" with confidence and credibility.


Curriculum Development Theory
System Theory
Adoption of Innovations Theory
Environment
Objectives
Social System
System/Environment
Structure
Innovation
Experience
Innovation/
Media Utilization
Process
Organization
Time
Transformation
Evaluation
Communication
Monitorial
System Evaluation/
Monitoring

Glassview of TPMDE


Curriculum Development
Systems Theory
Aoption of Innovations Theory
Environment
Objectives
Social System
Structural
Experiences
Innovation
Process
Organization
Time
Monitoring
Evaluation
Communication

"Original" TPMDE

Figure 1.

References

Baath, J. A. (1982). Teaching models for designing courses creatively. In J. S. Daniel, M. A. Stroud, & J. R. Thompson (Eds.), *Learning at a distance: A world perspective* (pp. 37-40). Edmonton, Alberta: International Council for Correspondence Education, Athabasca University.

Banathy, B. H. (1973). *Developing a systems view of education: The systems model approach.* Belmont, CA: Fearon Publishers.

Griew, S. (1982). The neglected role of disciplinary research in the distance university. In J. S. Daniel, M. A. Stroud, & J. R. Thompson (Eds.), *Learning at a distance: A world perspective* (pp. 188-191). Edmonton, Alberta: International Council for Correspondence Education, Athabasca University.

Holmberg, B. (1985). *The feasibility of a theory of teaching for distance education and a proposed theory.* ZIFF Papiere 60. FernUniversitat, Hagen, West Germany: Zentrales Institut fur Fernstudienforschung. (ERIC Document Reproduction Service No. ED 290 013)

Keegan, D. (1990). *The foundations of distance education* (2nd ed.). London: Croom Helm Ltd.

Rogers, E. M. (1983). *Diffusion of innovations* (3rd ed.). New York: The Free Press.

Shih, M. (1990). *Distance education in Taiwan: Applying the Triad Perspective Model of Distance Education to evaluate programs of the National Open University of Taiwan.* Doctoral dissertation, University of Northern Colorado, Greeley, CO.

Tyler, R. W. (1949). *Basic principles of curriculum and instruction.* Chicago: University of Chicago Press.

Wedemeyer, C. (1973). The use of correspondence education for post-secondary education. In A. Kabwasa and M. Kaunda (Eds.), *Correspondence education in Africa* (pp. 72-79). London: Routledge and Kegan Paul.

Wilson, R. W. (1970). *Learning to be Chinese: The political socialization of children in Taiwan.* Cambridge, MA: The M.I.T. Press.

Zvacek, S. M. (1990). Distance education in the teacher education program of Zimbabwe. Paper presentation for the International Division at the annual convention of AECT, Anaheim, CA, and published in *Proceedings of Selected Research Paper Presentations,* February, 1990.

Title:

Effects on Concretely versus Abstractly Illustrated Instruction on Learning Abstract Concepts

Authors:

Margaret Ann Smith
Patricia L. Smith

Abstract

Traditionally, illustrations used in instruction have been concrete in nature with the learning outcome not always considered in the design of the illustrations. This study investigated whether concrete or abstract illustrations were more effective for supporting acquisition of abstract concepts in print media when evaluated by immediate and delayed testing.

Introduction

The design of visuals for teaching concepts is an important issue for the instructional designer. Kirkpatrick as early as 1894 studied the retentional role of pictures though immediate and 72-hour delayed testing (Madigan, 1983). The results showed a positive superiority effect of pictures over words, especially in delayed testing. More recently, researchers have been evaluating the use of visuals in the teaching of concepts. The main focus of research has been on the teaching of concrete concepts with the use of concrete illustrations. However, this has left a gap in investigative research on the teaching of abstract concepts with the use of illustrations (Newby & Stepich, 1987).

Purpose

The purpose of this paper is to report the results of a study that investigated the question "What are the characteristics of effective instructional illustrations?" Specifically it investigated whether abstract or representational (concrete) illustrations were more effective in supporting the acquisition of abstract concepts.

Prediction 1

In the comparison of the illustrated versus the non-illustrated version of instructional text, it was expected that the illustrated version would be more effective than the non-illustrated version both on immediate and delayed tests. This effect was predicted due to two theoretical cognitive mechanisms of illustrations a) dual coding of illustrated information in the two representation systems of imagery and propositions (Paivio, 1986) and b) the attentional function of illustrations to the information represented within them (Duchastel, 1978; Peeck, 1987).

The dual coding theory suggests that even though the two subsystems work independently, one can have an effect on the other in terms of recall because of their partial interconnectedness. This implies that activity in one storage system can activate the other. If a bit of information which was stored both as an image and as a proposition cannot be retrieved from one storage area it may be retrieved from another, enhancing the retrievability of that piece of information. Based on this theory it can be predicted that information that is presented both as a image (whether concrete or abstract) and as text will be recalled better and longer than information that is presented only as text.

One primary mechanism for the advantage of illustrated over non illustrated text is the attentional function of illustrations. Also, much previous research on other types of learning such as verbal information and concrete concepts has supported the illustrated over non-illustrated versions of instruction (Dwyer, 1978; Levie & Lentz, 1982; Merrill & Tennyson, 1977). Consequently, it seems that these benefits will also support acquisition of abstract concepts.

Visual information can be retained over a much longer period of time than textual information (Paivio, 1979). Studies have shown that, when compared to immediate testing, delayed testing of materials results in more of the visual portion of a text being remembered than that of the verbal (eg., Levie & Lentz, 1982). This indicates that the processing for visual information seems to be different from the processing for verbal information, supporting the dual coding theory of separate paths and storage systems for verbal and visual information, as well as, the superiority of imaginal over textual long term recall. (Paivio, 1979).

Prediction 2

The authors predicted that on immediate testing over the content; the students from the concrete visuals group would score higher than the abstract visuals group. The concrete visuals group was given an illustration of a "best example" which represented the attributes of the concept in a form that could be quickly and easily remembered (Newby & Stepich 1987). The superiority of the concrete illustrations over the abstract visuals in the immediate test was predicted because it was expected that it would be easier for learners to remember the relationship of "best example" to the criterial attributes of the concept immediately after instruction, than to relate the attributes to the abstract visual.

Prediction 3

It was expected that on delayed testing the abstract visuals group would score higher than the concrete visuals group. This result was expected because it was believed that over time, the relationship of the criterial attributes to the concept class would be more memorable for students who received the abstract illustration than for those who received the concrete illustration. It was expected that an abstract embodiment of the abstract criterial attributes would provide a superior over-arching schema that would encourage better recall of the attributes of the concept over the long term than the concrete,"best example" illustration. According to theories of knowledge representation that conjecture a hierarchical storage system, such as Ausubel's (1963), a super-ordinate concept should cue recall of more nodes than an individual instance stored beneath it.

Research Questions

On an abstract concept learning task:

1. Will both immediate and delayed performance of those who receive abstract illustrations be superior to those who received no

illustrations and will those who received concrete be superior to those who received no illustrations?

2. Will immediate post test performance of those who receive concrete illustrations in instruction be superior to the performance of those who receive abstract illustrations?

3. Will delayed post test performance of those who receive abstract illustrations in instruction be superior to the performance of those who received concrete illustrations?

Participants

The participants in this study were junior and senior teacher education students enrolled in an instructional design and development course at a southwestern university (22,000 enrollment) in the summer of 1990. This class was selected because of the content relevance, availability of students and the typical print-based method of presenting content. All students in the classes were asked to volunteer for the study, and 68 agreed (because of attrition only 52 completed the study) to participate and were randomly assigned to one of the three study groups (no illustrations, concrete illustrations, or abstract illustrations). The only requirement for participants to be eligible for this study was enrollment in the instructional design and development course and availability for delayed testing.

Demographic data/entry questionnaire was collected at the beginning of the research study through a survey given before the lesson. Participants fitted into the following categories: 14 males, 37 females (one person evidently did not know what gender s/he was!), 3 sophomores, 18 juniors, 21 seniors, 10 graduate students. The most common majors were, elementary education, early childhood education and secondary social studies education. The average participant age was 23, with ages ranging from 19 to 39. Forty-six percent of the students had had courses in learning theory. Seventy-five percent reported previous experience with Bloom's (1956),taxonomy and 69% with Gagne's (1984), types of learning outcomes. Yet, when asked what Gagne's types of learning outcomes were, only four percent remembered anything about them. It was determined from this survey that none of the students possessed sufficiently detailed prior knowledge of the content to be eliminated from the study.

Materials

The researcher-developed materials used for this study are three versions of a print-based lesson teaching the four categories within the intellectual skills domain of Gagne's domains of learning outcomes. The topic was chosen because of the abstract nature of the content. Also the topic filled the need for a subject that would be of value to the participants in the study, but did not assume much prerequisite knowledge or skill, thus increasing the motivation to learn the content. The instructional objective for the materials was: "Given a description of a learning task, the learner will be able to identify which of the intellectual skills (i.e., discrimination, concept, rule or problem solving) it represents." The abstract illustration treatment (A) contained instruction with illustrations which abstractly represented each of the four classifications. For example, the concept "rule/principle" was abstractly represented by circles, arrows and a rectangle representing relational rules (see Figure 1). The concrete illustration treatment (C) includes illustrations representing "the best example" of the concept. For example, the concept "rule/principle" was concretely represented by a visual of balloons expanding in a hot car, illustrating the rule "a gas, when heated, expands in volume," (see Figure 2). The no illustration treatment (B) was instruction with no visual illustrations.

Figure 1

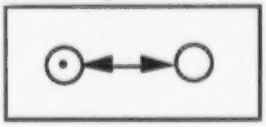
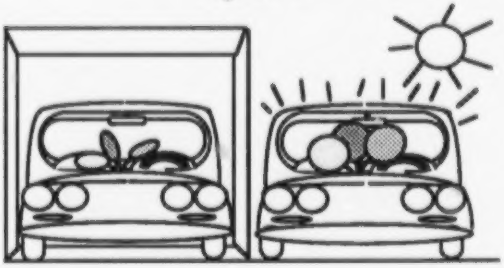


Figure 2



All three treatments have the best example explained verbally in the text as part of the lesson, due to the powerful effect of prototypes in concept learning (Newby & Stepich, 1987). Each set of materials for the study contained the same text content with the exception of sentences (two or three) that specifically referred to the illustrations, as traditionally recommended by Brody and Legenza, (1980). These caption sentences were constructed to convey no new content information.

The lesson design followed Gagne's nine events of instruction. It also incorporated Merrill and Tennyson's (1977) instructional design guidelines for teaching concepts. The Gagne's nine events are numbered and listed below along with the explanation of how each one was achieved in the stimulus lesson:

1. Gaining attention: The fact that this was not a normal classroom situation helped get the attention of the students. The researcher verbally explained the importance of the lesson content to the class and the instructions to be followed. Attention was also gained

in the initial textual references to the relevancy of the instruction to teaching tasks.

2. Informing the learner of the lesson objective was presented verbally in the first paragraphs of the materials.

3. Stimulating recall of prior learning: In the first paragraph of the lesson a reference to prior learning was made. "From your experiences in learning in all your years of public education as well as in the University you have probably noticed that there is a difference among learning tasks that have been assigned. Not only are some tasks harder, some take more time, some use different types of learning techniques."

4. Presenting the stimulus material with distinctive features: The lesson presented each distinctive feature (criterial attributes) of each of the learning outcomes (discriminations, concepts, rules and problem solving) and examples of each.

5. Providing learning guidance: Sentences were included to cue learners to effective learning strategies, such as: "use this example to help you remember...".

6. Eliciting performance: At the end of each section there was a review of the concepts being learned:

> "Practice:
> Please circle the numbers of the following examples which are evaluating learning of rules.
> 3. Which pronoun is correct? "The dinner club was too expensive for Jim and me/I".

In addition there is a cumulative practice at end of the lesson.

7. Providing informative feedback:
Informative feedback was placed at the end of each of the practice sections and at the end of the cumulation practice section, for example, "Questions 2, 3 and 4 are examples of concepts because students must classify according to characteristics and identify by the names 'propaganda,' 'Datsun,' and 'barn.'"

8. Assessing performance: Learning was assessed in the post test and the delayed post test.

9. Enhancing retention and learning transfer: The opportunity for learning transfer and reinforcement comes in the lesson, practice, review, post test and delayed post test. The post questionnaire and oral interview may also have been helpful in requiring the students to think about the use of the visuals in remembering the content of the lesson and how to use them in the future with formerly unencountered situations.

Tests

The testing instruments, both immediate post (range 0-30) and delayed post (range 0-31), for the instructional materials consisted of parallel forms of questions assessing classification of examples of each of the four learning tasks. The tests were over the textual content only.

The illustrations were expected to assist the students in recalling the information but there was no reference to the illustrations on the tests nor any question that could not be answered from the textual material. Similar questions (e.g. based on the same objectives) were given on both the immediate recall and delayed tests.

The first 24 test items on each test evaluated near transfer of the instructional objective by asking the students to classify concepts by selecting which one of the four intellectual skills that the statement represented. The following is an example of the test items: "How do you locate and fix the source of a problem of no sound in an audio system?" This intellectual skill is problem solving. The two tests used were parallel forms with equal number of each of the items testing conceptual knowledge of four concepts (four types of skill) on each of the two tests. The last part of the immediate post and delayed post test were six or seven far transfer items on each test, respectively. These items included making predictions and inferences based on concepts, explaining answers on categorization examples, explaining distinguishing characteristics in concept examples, supplying their own examples and explaining relationships among the related concepts with such questions as:

Which of these is the best strategy for learning the rule on use of pronouns? (Tom gave the bike to Bill and me.)

____a. Making flash cards with written statement of the grammar rule.

____b. Explaining the rule in written form.

____c. Developing your own examples of the use of the rule.

____d. Writing the rule 20 times.

The immediate post test score was chosen as a dependent variable to test information immediately after the lesson. The time period of one week was a logical choice for delayed testing because the information being tested would normally need to be stored in memory for about one week for use in the following class.

A post questionnaire using a Likert scale and completion questions were given after the immediate post test to determine attitudes and perceived understanding of the lesson and the illustrations. In addition, a 10% sample of the student participants from each treatment group was interviewed orally regarding their understanding and use of the illustrations between the immediate and delayed post-tests.

All materials developed by the researcher were read for validation of content by three instructional design experts in the educational technology area. Post test and delayed post test data from the study was used to obtain a reliability test of instrument at .70 and .66, immediate and delayed, respectively.

The test was scored by a key for the first 24 questions of each test and according to preset criteria for the last part of the test. If the answer did not meet the criteria for that question it was counted incorrect. There was no partial credit for a question.

Procedures

The class roll was used to assign the experimental treatments to the students by assigning the first person on the roll to the non-illustrated version, the second to the concretely illustrated version and the third person to the abstractly illustrated version, and so on.

Data collection occurred on three separate days. The first day of the regular class period of the semester consisted of the completion of an agreement to participate (strictly voluntary) form and the collection of demographic data/entry questionnaire for use in reporting information about the participants and any prior knowledge of the lesson content. On the second meeting day the instructions and the purpose for the lesson were explained before the lesson started. The lesson was placed in a brown envelope with the person's name and the last four digits of his/her social security number on the outside. The materials inside were given a unique identification number. The materials were color coded by treatment type and by lesson versus test to help with the instructions. When the students completed the lesson they replaced it in the envelope before removing the immediate post test. When the posttest was completed the students filled out the post questionnaire, placed all the material back in the envelope and left.

The unannounced delayed posttest was given one week after the immediate posttest. The delayed test was placed in the same envelopes in which the lesson and immediate posttest were placed and assigned the same packet's unique number to pair up other data with the posttest results.

A ten percent sample of the population was interviewed after the immediate posttest to gain more information on the way in which the participants used/did not use the visuals in the lesson. The questions asked in the oral interview were taped for accuracy in reporting. The questions consisted of the students' use and need for the visuals in remembering the lesson content. A sample question was "Did the illustrations help you remember the information?"

Variables

Independent variable

The independent variable in this test was type of illustration with three levels: none, concrete and abstract. These variables were embodied in three print-based instructional versions of a lesson on one of Gagne's domains of learning outcomes, intellectual skills. Each lesson was identical except for the presence of abstract illustrations, concrete "best example" illustrations or no illustrations. Both illustrated versions had the same pattern of layout of verbal information on two thirds of the page on the right side and visual information or white space on the left one third of the page. Textual information referring to the illustrations was placed in the same pattern on each illustrated version.

Dependent Variable

The dependent variables in this study were immediate and delayed post test scores discussed under the test section.

Design

The design of this study was a immediate-test-post-test control group design. Data was analyzed with a one-by-three way factorial analysis of variance (ANOVA). The illustration variable had three levels: none, concrete and abstract. The dependent variables were immediate and delayed post-test scores. This design is shown in Figure 3.

Figure 3

Design

	abstract	Illustration Factor not present	concrete
Immediate testing	**N=17**	**N=16**	**N=19**

	abstract	Illustration Factor not present	concrete
Delayed testing	**N=17**	**N=16**	**N=19**

Results

The means and standard deviations were computed for each group as shown in Table 1 for the immediate post test and Table 2 for the delayed post test. A 1 X 3 factorial analysis of variance for the immediate post test resulted in (F=2.0, df=2, 49, p=.15) and for the delayed post test (F=0.45, df=2, 49, p=.64). The alpha level was be set at the .05 level of significance. The results were not statistically significant.

Table 1
Immediate Post Test

Group	N	Mean	Standard deviation	range
A	17	21.65	4.21	11-28
B	16	21.44	4.2	11-28
C	19	23.68	2.71	20-28

Table 2
Delayed Post Test

Group	N	Mean	Standard deviation	range
A	17	23.24	3.73	14-30
B	16	22.63	3.88	14-28
C	19	23.79	3.31	18-29

Prediction 1 stated that the illustrated versions would be more effective than the non-illustrated version on both immediate and delayed tests. The results showed that the participants receiving the abstract and the concrete versions of the instructional lesson did better on both the immediate and delayed tests than participants receiving the non-illustrated version, although the actual statistical comparison of these data was not computed since the overall F was not significant.

Prediction 2 stated that on immediate testing over the content; the students from the concrete visuals group would score higher than the abstract visuals group. The results showed that the participants receiving the concrete illustrations did tend to score higher on the immediate tests than the participants receiving the abstract illustrations. However, this comparison of these data was not computed since the overall F was not significant.

Prediction 3 stated that it is expected on delayed testing the abstract visuals group will score higher than the concrete visuals group. The results showed that the participants receiving the abstract illustrations scored lower than the participants receiving the concrete illustrations.

The data were subjected to a posthoc analysis to analyze the gain scores from immediate post to delayed post, which showed the abstract groups' mean difference tended to gain more over time than both the concrete and non-illustrated groups. Also the non-illustrated group tended to gain more than the concretely illustrated group. These results were not statistically significant to the .05 level (F = .36, df = 2,48, p = .7).

Figure 4

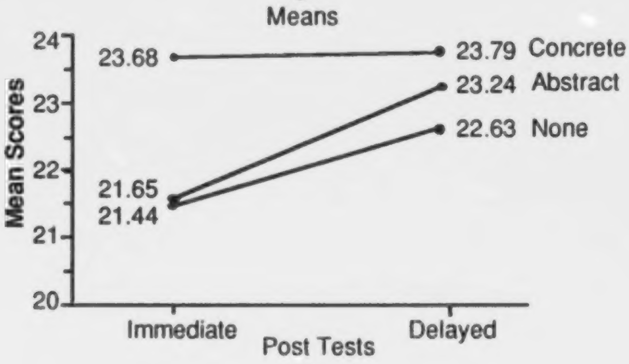


Figure 4 shows a line graph of the means of immediate and delayed post test scores over a one week period. Table 3 shows the means of the gain between the immediate and delayed for the three groups are: 1.59 for the abstract group, 1.19 for the no illustration group and 0.11 for the concrete illustration group.

Table 3
Means and Gain

	Abstract Illustration	No Illustration	Concrete Illustration
	N=17	N=16	N=19
Tests	Means	Means	Means
Immediate Post	21.65	21.44	23.68
Delayed Post	23.24	22.63	23.79
Gain from I P to D P	1.59	1.19	0.11

Discussion

This study investigated the comparative effects of concretely illustrated instruction versus abstractly illustrated instruction on the acquisition of abstract concepts. "On an abstract concept learning task will immediate post test performance of those who receive concrete illustrations in instruction be superior to the performance of those who receive abstract illustrations?" Although the data tended to support the conclusion that concrete > abstract on immediate post-test, these differences were not statistically significant. These findings somewhat support the findings of prototype "best example research". In addition, although abstract illustrations < concrete on delayed testing , it was indeed interesting, that the abstract visuals tended to increase retention to a greater degree than the concrete visuals as shown by the gain from Immediate posttest to delayed post (see Table 3). This finding will be followed up in later studies.

There are several factors that may have limited the robustness of this study. The lesson was written in such a way that was not difficult with or without visuals. For instance the non-illustrated group's mean performance was 22.3 (out of 30) which was quit high. The abstract visuals may have been too detailed or the verbal description of them not detailed enough. Because of the preliminary evaluation of the materials the researchers came to the conclusion of the need for further evaluation of both the test content and abstract and concrete visuals used.

In conclusion many factors have been found to affect learning with the use of illustrations in instructional text: the learning outcomes desired, the methods of presentation, the type of verbal reference to the illustration with captions and/or labels and/or direct reference within the text, how text relevant the illustration is and how dependent the content is on the illustration. These factors should be controlled in further research with the learner's characteristics, such as reading ability and visual literacy, in mind. Much research on illustrations is still needed to sort out all the factors involved in the use of illustrations in instructional text. No matter what style of illustration is chosen, insuring that the illustrations provide the conditions needed to support each particular type of learning seems to be the deciding factor in designing illustrated instructional text.

References

Ausubel, D. P. (1963). Psychology of Meaningful Verbal Learning. New York: Grune and Stratton.

Bloom, B. S., M. D. Englehart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. (1956). Taxonomy of Educational Objectives: Handbook I, Cognitive Domain. New York: McKay.

Brody, P. J., & Legenza, A. (1980). Can pictorial attributes serve mathemagenic functions? Educational Communication and Technology Journal, 28, 25-29.

Clark, R. E. & Salomon, G. (1979). Media in Teaching. In M. C. Wittrock (Ed.), Handbook of Research on Teaching: 3rd edition. New York: MacMillan Publishing Company.

Duchastel, P. C. (1978). Illustrating instructional texts. Educational Technology, 18(11), 36-39.

Dwyer, F. M. (Ed.)(1987). Enhancing Visualized Instruction-Recommendations for Practitioners. PN: Learning Services.

Dwyer, F. M. (1978). Strategies for Improving Visual Learning. PN: Learning Services.

Gagné, R. M. (1984). The Conditions of Learning and Theory of Instruction (4th ed.). New York: Holt, Rinehart and Winston.

Levie, W. H., & Lentz, R. (1982). Effects of text illustrations: A review of research. Educational Communications Technology Journal, 30(4), 195-232.

Madigan, S. (1983). Picture Memory. In J. C. Yuille (Ed.), Imagery, memory and cognition, Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers

Merrill, M. D. & Tennyson, R. D. (1977). Teaching Concepts: An Instructional Design Guide. Washington, D. C.: Educational Technology Publications

Newby, T. J. & Stepich, D. A. (1987). Learning Abstract Concepts: The Use of Analogies as a Mediational Strategy. Journal of Instructional Development, 10(2) 20-26.

Paivio, A. (1986). Mental Representations. New York, NY: Oxford University Press, Inc.

Paivio, A. (1979). Imagery and Verbal Processes. Hillsdale, NJ: Lawrence Erlbaum Associates.

Peeck, J. (1987). The Role of Illustrations in Processing and Remembering Illustrated Test. In D. B Willows & H. A. Houghton (Eds.), The Psychology of Illustration: Vol.1. Basic Research. New York: Springer-Verlag.

Title:

The Validity of a Multiple-Choice, Paper and Pencil Instrument in Discriminating between Masters and Nonmasters of Instructional Design

Authors:

Sidney Leland Stepp
Sharon A. Shrock

The Validity of a Multiple-Choice, Paper and Pencil Instrument in Discriminating Between Masters and Nonmasters of Instructional Design

Instructional design is a big business in industrial organizations. As the instructional design profession continues to grow and greater investments of time and money are made, instructional designers may be held more and more accountable for the instructional decisions made and programs developed. Questions of ability or selection will become more prominent and instruments that can validly be used in assisting to make judgements will become a paramount issue.

Assessing competency in instructional design has been a hot topic for debate for a number of years. While surveys have identified instructional design competencies, there has been virtually no systematic research of alternative means for assessing professional competence. The purpose of this paper is to report a study investigating the question: Can a multiple-choice, paper and pencil test validly discriminate between masters and nonmasters of instructional design?

The Issues Involved

In considering the study two sets of issues emerge: psychometric and political. The psychometric issues involve test item formats and the nature of the instrument. Seven item formats were considered: true/false, matching, fill-in-the-blank, short answer, essay, multiple-choice, and performance in an assessment center. Of the seven, actual performance in an assessment center is the most valid form of identifying instructional design competencies although assessment centers would be expensive in terms of the time required to perform the assessments and money required to establish and implement the appropriate testing. These problems make the assessment center approach impractical in most circumstances. Other item formats are inappropriate because of the inability of the formats to get at higher cognitive levels or difficulties in achieving scoring reliability. A multiple-choice instrument would overcome the problem of expense, provided that a valid, discriminating instrument could be developed.

The choice between a norm- and criterion-referenced instrument is difficult. Politically, the field of instructional design would most readily accept a criterion-referenced instrument or at least a norm-referenced instrument that covers all of the competencies found important for an instructional designer. Unfortunately, this is not possible with this type of instrument. The instrument cannot cover all of the competencies such as consulting skills or writing ability. A statistically validated, norm-referenced instrument was selected as the best choice for a multiple-choice instrument.

Political issues involve the different focuses between instructional design organizations and professionals. For more than the past 20 years, the issue of instructional design testing and certification has been problematic at both the organizational and individual levels. While organizations express an interest in the certification issue and competency testing (Prigge, 1974), no progress and few efforts have been made as demonstrated in the literature. Organizations have not united to work together perhaps due to their varying special interests and audiences. In the single case where people actually sat down together to develop a criterion-referenced test with objectives-based test items (NSPI/DID-AECT Task Force in 1982), the varying backgrounds of the group provided an overwhelming stumbling block (Sharon A. Shrock, personal communication, May 10, 1989). Failed attempts aside, the research has not addressed the question: Can valid items be constructed that can make fine discriminations for an instructional design competency instrument?

The Three Stages of Instrument Development

The instructional design instrument was developed in three stages illustrated in Figure 1. First, items were composed and revised until subject matter experts agreed to each item's logical validity. Second, trial testing and item analysis were done to test empirically and eliminate non-discriminating items from the instrument. In this stage two groups of subjects contributed to the data: one group of non-professional instructional design masters and one group of instructional design nonmasters. In the third stage a phi coefficient was calculated to show a level of concurrent validity. The Tukey method of multiple range means testing was used to show significant differences between groups and subgroups. Multiple discriminant function analysis was used to identify additional items that did not serve to discriminate between groups. In this third stage two new groups contributed to the data: professional masters and nonmasters. In addition, the nonmasters group was divided into four subgroups: education graduate students, education undergraduate students, non-education graduate students and non-education undergraduate students. The nonmasters were split into these four subgroups to obtain information concerning the instrument's ability to make fine discriminations between overlapping groups.

Composing and Revising Items

Item composition began following a brainstorming session with subject matter experts. At that meeting the specific means of item writing and stages of item analysis were discussed. Several conclusions were reached.

Resources for the Item Bank

The multiple-choice items for the item bank were written. Non-knowledge level principles and concepts (as classified by Bloom's and Gagne's taxonomy) used for the items were selected from instructional design textbooks listed in Appendix A.

The 50 item bank and subject matter expert review. The items were reviewed by subject matter experts. During the meetings, the subject matter experts discussed item clarity, ambiguity, and logical validity. Suggestions for item changes or removal were made. A total of 35 items were agreed upon for the instrument's item analysis.

Trial Testing and Item Analysis

Groups of masters and nonmasters of instructional design were identified by the subject matter experts. After volunteering to participate, the master and nonmaster groups were asked to complete a demographic data sheet and the instrument.

Non-Professional Masters

A total of 17 current and former graduate students in the Department of Curriculum and Instruction at Southern Illinois University were identified by the subject matter experts as having a mastery of instructional design. The mastery decision was based on the subject matter experts' observations of the students' course work in instructional design, work with clients, and interactions in the classroom. Of this group, 16 students completed the instrument.

Nonmasters

A total of 57 nonmasters of instructional design were identified by the subject matter experts. The selection of the nonmasters from these particular classes was based on the subject matter experts' knowledge of the students and their abilities (or inabilities) in instructional design rather than the courses taken by the students in instructional design.

Item Analysis

Group means for the 73 non-professional masters and nonmasters are illustrated in Figure 2. Pearson point-biserial coefficients ranged from -.14 to .41 for the instrument items. Seven items were removed from the instrument that had negative Pearson point-biserial coefficients. Item analysis was performed again using the data from the 28 remaining items. Pearson point-biserial coefficients ranged from .09 to .54 with no negative Pearson point-biserial coefficients remaining in the data set. The Cronbach alpha coefficient increased from .5521 for the 35 item instrument to .6574 for the 28 item instrument.

Three other item analyses were performed removing other items with low Pearson point-biserial coefficients to find if results could be further improved. No other analysis provided better results than the initial 28 item analysis. The 28 item instrument was selected for further data collection and instrument validation.

Concurrent Validity

At this point in the research new data was collected with the 28 item instrument for instrument validation. To enhance the generalizability of the data analysis results it was important to broaden the population used to include subjects outside the College of Education at Southern Illinois University and outside the Southern Illinois University environment. New groups of masters and nonmasters of instructional design were identified and solicited to complete the instrument.

A total of 48 professional masters of instructional design concepts were identified from the membership of the National Society for Performance and Instruction and the Association for Educational Communications and Technology. Packages containing the instrument materials, an addressed, stamped envelope, and a letter briefly describing the research were sent to the 48 professional masters.

Professional Masters

The 48 professional masters (PM) were selected for their known expertise in the field of instructional design. Geographically, the group was spread across the United States and Canada. A total of 34 completed instruments and three incomplete instruments were returned providing a total response rate of more than 77%.

Demographic data. While 34 completed instruments were returned, demographic data sheets were not returned for two of the instruments. The following data reflect the 32 demographic data sheets that were returned. Also, some subjects did not complete all items on the demographic data sheets. The following data reflect completed items for the sheets that were returned.

Table 1 shows a summary of the demographic data concerning the respondents' jobs collected from the professional masters. As can be seen from the table, respondents stated that they held one of six types of positions ranging from academic faculty to business executive to private consultant. Respondents were equally split between academic and business affiliation with 16 respondents from each of the two types of positions. Table 2 shows the degrees earned by the professional master respondents. The majority (29) responded that they had a Ph.D. or an Ed.D. while 1 held an M.A. and 1 held a B.A. All but one subject responded that jobs held were instructional design related. A total of 28 respondents indicated that positions held were related to education while 3 indicated that positions held were not educationally oriented. In all three of these cases the job held was corporate management.

Table 3 shows the programs where the degrees were obtained by the professional masters. Eleven different programs were indicated. The largest number of respondents indicated that they obtained their degrees in programs called instructional systems technology. Additionally, not all programs were directly related to instructional design such as anthropology and psychology.

Finally, the range of hours of course work in instructional design taken by the professional masters varied greatly. A total of 19 respondents indicated that they had taken more than 12 hours, 3 respondents indicated 6 to 12 hours, 1 respondent indicated 3-6 hours, and 8 respondents indicated less than 3 hours. The relatively large number of subjects responding that they had taken less than 3 hours could be due to the newness of instructional design programs combined with the ages of the respondents (i.e., there are three types of people in the field of instructional design: experience without instructional design education, instructional design education without on-the-job experience, and experience with instructional design education.) One respondent wrote a note indicating that the respondent's degree was obtained before courses in instructional design were offered in the respondent's program. Respondents were asked to complete the instrument because of their known expertise in the field. In some cases this expertise was based on years of work experience. Some respondents had not attended recently developed programs of instructional design to learn theories about what they were already practicing and in many cases publishing in the field.

Nonmasters

The nonmasters group was composed of students enrolled at Southern Illinois University. The group contained four subgroups: 43 education graduate students (EGS), 45 education undergraduate students (EUS), 23 non-education graduate students (NEGS), and 39 non-education undergraduate students (NEUS).

Descriptive Analysis

Table 4 shows the descriptive analysis of the instrument responses for the 184 professional masters and nonmasters in the study. The table shows the professional master group and nonmaster subgroup, ranges, means, and standard deviations. As can be seen in the table, the professional master group and four nonmaster subgroup numbers range from 23 to 45, the means range from 7.949 to 19.265, and the standard deviations range from 2.516 to 3.336. Figure 3 illustrates the various mean scores for the professional master group and the four nonmaster subgroups.

Setting the Cut Off Level and Demonstrating Concurrent Validity

Figure 4 shows the smoothed frequency distributions of scores for the professional masters overlaid with the frequency distribution of scores for the nonmasters. The figure shows that the distributions for the two groups intersect at a score of 17. As described by Allen and Yen (1979) this intersection of frequencies can be accepted as the cut off level for computing phi to demonstrate the validity of the instrument. The consequences of choosing a score of 17 as the cut off level are two-fold. First, some identified masters will be misclassified by the instrument as nonmasters. Second, some identified nonmasters will be misclassified as masters by the instrument. The main objective in setting this cut off score is to minimize the misclassification for both consequences.

A phi correlation was used to demonstrate the instrument's concurrent validity. Figure 5 shows the arrangement of data used for the calculation. The figure also shows the percent of the master/nonmaster groups falling into each of the four categories. The four categories are formed from the master/nonmaster classification of the subject matter experts and the master/nonmaster classification of the instrument at a cut off level of 17. The figure shows that the instrument would misclassify 4 masters or 11.76% of the master group and 16 nonmasters or 10.67% of the nonmaster group.

At a cut off level of 17 the phi coefficient produced was .695. Subsequent runs with other cut-off levels (14, 15, 16, 18, 19, and 20) did not produce a greater phi coefficient demonstrating the validity of the choice of 17 as the cut off level. At a cut off level of 16 a phi coefficient of .601 was produced. At this level approximately 6% of the masters would be misclassified and approximately 21% of the nonmasters would be misclassified. At a cut off level of 18 a phi coefficient of .626 was produced. At this level approximately 7% of the nonmasters and about 29% of the masters would be misclassified.

Tukey Method of Multiple Range Means Testing

Table 5 shows the results from running the Tukey Method of multiple range testing of the five groups/subgroups of means. As is shown in that table, using a harmonic N ($\underline{N} = 34.72$) to obtain an average sample size, the Tukey method would require a 1.82 difference between mean scores for significance at $\underline{p} < .05$. The table shows that the mean score of the professional masters was significantly higher than the mean score of students in education, the mean score of students in education was significantly higher than the mean score of graduate students not in education, and the mean score of graduate students not in education was significantly higher than undergraduate students not in education. What is surprising in the results is that the mean scores of graduate and undergraduate students in education were *not* significantly different. At the same time the mean score of undergraduates not in education was significantly lower than the graduate students not in education. Differences between mean scores do not appear to exist because of differences between graduate and undergraduate abilities.

A Look at the Individual Items and Multiple Discriminant Function Analysis

Multiple discriminant function analysis was performed using the 28 instrument items as group membership predictors to identify any items that did not predict group membership. The five groups used in the analysis were professional instructional designers (PM), education graduate students (EGS), education undergraduate students (EUS), non-education graduate students (NEGS), and non-education undergraduate students (NEUS). The univariate F tests demonstrated that all but four of the items significantly discriminated between groups, $\underline{p} \leq 0.05$. The Wilk's Lambda, also computed in the analysis, demonstrated similar results to the Tukey multiple range analysis. The groups do differ significantly (aside from the education graduates and education undergraduates) on all instrument items as a set. The Wilks' lambda was calculated to be .063. This is equivalent to a statistically significant $\underline{F}(112, 606) = 5.458$, $\underline{p} \leq .05$.

On the basis of all 28 predictors, Chi square tests were computed for each of the four derived discrimination functions (based on five groups minus one) to determine the significance of discrimination along each dimension. The first discriminant function was found to be significant $\underline{\chi}^2(112, \underline{N} = 184) = 461.09$, $p \leq .05$. The second and third functions were also found to be significant $\underline{\chi}^2 (81, \underline{N} = 184) = 200.58$, $\underline{p} \leq .05$ and $\underline{\chi}^2(52, \underline{N} = 184) = 96.149$, $\underline{p} \leq .05$. The fourth dimension failed to reach the necessary level for significance (p = .30). The first three discrimination functions accounted for 97% of the variance between groups. Greater coefficient values show a greater ability of the item to discriminate between groups. Four items on the instrument had low values. The univariate analysis indicated that these items did not help the instrument discriminate between the five groups.

Figure 6 shows the discriminant analysis results for the predicted versus actual classification for the total sample. The diagonal line in the figure underlines those numbers of correct classifications. The high number of correct classifications demonstrates the ability of the whole instrument to discriminate between groups.

New Phi Correlation After Removing Four Items

The four items that discriminant analysis identified as not useful in discriminating between groups were removed in order to recalculate a new phi coefficient using only the items that did statistically work to discriminate between groups. New professional master group and nonmaster subgroup mean scores are shown in Figure 7. All group and subgroup mean scores are lower than the means shown in Figure 3 although it must be remembered that the mean scores shown in Figure 7 are based on a 24 item instrument while those in Figure 3 are based on a 28 item instrument. The phi coefficient after removing the four items increased from .695 to .758. The Cronbach alpha coefficient also increased from .746 to .762. Removing the four items statistically increased the validity and reliability of the instrument.

Conclusions

A total of 257 subjects completed the instrument. Item analysis and instrument validation were performed on the data. Several conclusions can now be drawn concerning the instrument, its validation, and its future.

The Validity of the Instrument

A test of concurrent validity compares mastery classifications. A comparison between the mastery classification of the subject matter experts and the classification of the instrument at a cut off level of 17 produced a phi coefficient of .695. The instrument's concurrent validity in this study has been established.

Univariate F tests in a discriminant analysis. Univariate F tests during discriminant analysis of the instrument items showed that 4 items did not serve to significantly discriminate between groups. The removal of the 4 items increased the phi coefficient produced by the data from .695 to .758. The Cronbach alpha coefficient was also increased from .746 to .762 after the removal of the 4 items. The instrument can be further refined in future studies.

Wilks' lambda in a discriminant analysis. Before discriminant functions could be generated, the five groups of data needed to be tested to see if they differed significantly on the 28 instrument items as measured by the Wilk's lambda statistic. The Wilks' lambda was calculated to be .063. This is equivalent to an F (112,606) of 5.458, $p \leq .05$. The instrument items do discriminate among the five groups.

Discriminant analysis. Chi square tests were computed for the derived discrimination functions to determine the significance of discrimination along each of the four dimensions. The first three discriminant functions were found to be significant (χ^2(112, N = 184) = 461.09, $p \leq .05$; χ^2(81, N = 184) = 200.58, $p \leq .05$; and χ^2(52, N = 184) = 96.149, $p \leq .05$), but the significance of the fourth dimension failed to reach the necessary level (p = .30).

A Comparison of Professional and Nonprofessional Masters' Mean Scores

The mean scores of the non-professional masters used in the item analysis stage and the professional masters used in the validation stage are illustrated in Figure 8. The mean score of the professional masters ($\underline{M}$ = 19.265) was not significantly different from the nonprofessional masters ($\underline{M}$ = 19.5). In view of the years of real world experience of the professional masters and the inexperience of the nonprofessional masters this non-significant result might seem strange. Experience would seem to add to the ability of the professional masters over a simple knowledge of the theories studied by the nonprofessionals without the real world experiences. Since the profession of instructional design is relatively new and studies comparing the knowledge bases of experienced professionals and non-experienced professionals do not exist, an analogy to the medical profession (where studies of this nature have been performed) seems appropriate.

A key difference found in studies comparing newly graduated medical students with experienced doctors is time. When time in making a decision is not a factor, "student recall will exceed experts" (Schmidt, Norman, & Boshuizen, 1989, p. 17). On the other hand, when time becomes a factor, "the trend reverse[s] and experts recalled more than novices (Schmidt, Norman, & Boshuizen, 1989, p. 17). An explanation of these phenomena is provided by Norman (1990) "since expert knowledge is compiled and [newly graduated students] are actively elaborating mechanisms, [newly graduated students] will recall more, but will require more time to process the text. Thus under conditions of unrestricted time. . .student recall will exceed experts." Other studies comparing clinical experience and expertise show no differences between experienced experts and newly graduated students (Feltovich, Johnson, Moller, & Swanson, 1984). Again, the key factor is time. Experienced professionals have the situations that they have seen in the past to act as templates for new situations that they see in the present. For example, a man comes to an experienced doctor with symptoms of vomiting and intestinal cramps. At the same time, the doctor notices that the man's skin has a yellow tinge. An experienced doctor might be able to relate the case to a similar set of symptoms from a person treated last month, last year, etc. An inexperienced doctor would need to start from scratch putting all of the symptoms together to diagnose the illness taking more time than the experienced doctor. Given time restrictions, an experienced professional performs better than an inexperienced professional in medicine.

The instrument in this study was used without time restrictions. If the medical explanation of experienced versus inexperienced differences is an applicable explanation for the field of instructional design, the time factor could be one explanation for the non-significant difference between mean scores of the professional masters and the non-professional masters.

Recommendations and Summary

At this point we would like to make recommendations regarding the use of this instrument and the future research of an instrument of this type.

A Research Tool

While the idea for the instrument came about from a continuing dialogue concerning certification in instructional design, it was and is not expected that this instrument be used in such a process. It is a research tool for use in a person's lifelong research agenda.

Further Research

Further research is needed for two reasons. First, many of the professional masters wrote helpful notes and suggestions for further item refinement as they completed items. It is felt from those responses that some changes in the instrument need to be made. Suggestions included grammatical changes to enhance question clarity and item content changes to reduce ambiguity. The second reason for further research is to broaden the scope of the subject groups. In the current study nonmaster subject groups were primarily students in education and science. Using the instrument with other professional groups would provide more information about the instrument's discriminating abilities. Other professional groups might include business administrators and professional trainers. Because of the overlapping competencies between those fields and instructional design, questions concerning how finely the instrument can discriminate could be addressed.

A paper and pencil, multiple-choice instrument can validly discriminate between masters and nonmasters of instructional design although further research is needed. The field of instructional design is quickly growing. As the field continues to grow and gain in importance, instructional designers will be expected to be more accountable for their abilities and actions. Questions of ability will become more prominent and questions such as the one in this study will need to be answered along with this higher demand for accountability.

Table 1
Professional Masters Demographic Data--Jobs Held

Job Held	Response Frequency
Corporate Educational Specialist	1
Corporate Executive	9
Corporate Instructional Developer	3
Faculty	16
Human Factors Specialist	1
Private Consultant	2

N = 32

Table 2
Professional Masters Demographic Data--Degrees Held

Degree	Response Frequency
Ph.D.	24
Ed.D.	5
M.A.	1
B.A.	1

N = 31

Table 3
Professional Masters Demographic Data--Program or Department in Which Degree Was Obtained

Program or Department	Response Frequency
Anthropology	1
Curriculum and Administration	1
Curriculum and Instruction	3
Education	2
Educational Psychology	1
Educational Technology	2
Instructional Design	2
Instructional Systems Technology	10
Instructional Technology	4
Psychology	4

N = 30

Table 4
Descriptive Statistics

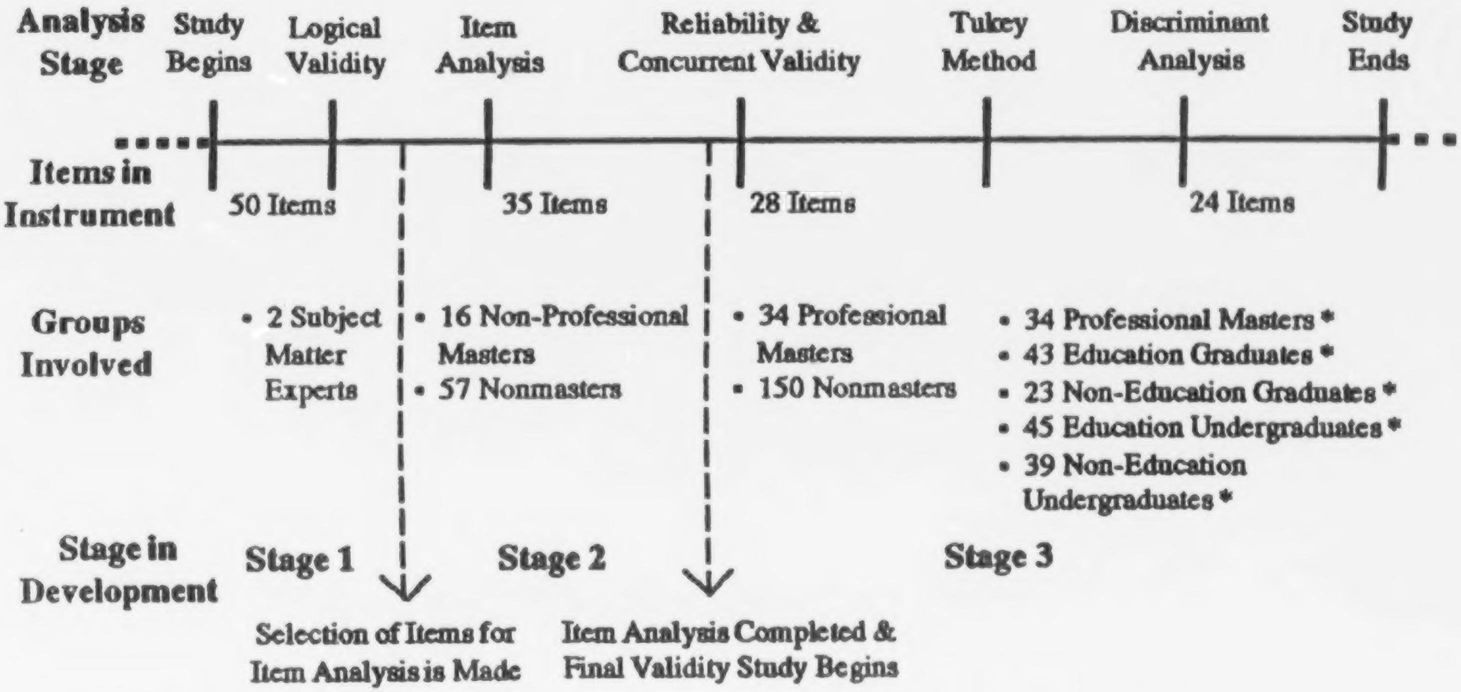
Group	n	Maximum	Minimum	Mean	Standard Deviation
All Subjects	184	25	3	13.310	4.601
Professional Masters	34	25	14	19.265	2.632
Education Graduates	43	19	5	14.535	2.881
Non-Education Graduates	23	16	3	10.696	3.336
Education Undergraduates	45	19	9	13.622	2.516
Non-Education Undergraduates	39	13	4	7.949	2.523

Table 5

Tukey HSD Multiple Comparison Test for the Professional Master Group and the Four Nonmaster Subgroups

Group	Professional Masters	Education Grad.	Education Undergrad.	Non-Education Grad.	Non-Education Undergrad.
	N = 34	N = 43	N = 45	N = 23	N = 39
Mean	19.265	14.535	13.622	10.696	7.949

p ≤ .05
Harmonic N = 34.720
Critical Difference = 1.813


Analysis Stage
Study Begins
Logical Validity
Item Analysis
Reliability & Concurrent Validity
Tukey Method
Discriminant Analysis
Study Ends
Items in Instrument
50 Items
35 Items
28 Items
24 Items
Groups Involved
• 2 Subject Matter Experts
• 16 Non-Professional Masters
• 57 Nonmasters
• 34 Professional Masters
• 150 Nonmasters
• 34 Professional Masters *
• 43 Education Graduates *
• 23 Non-Education Graduates *
• 45 Education Undergraduates *
• 39 Non-Education Undergraduates *
Stage in Development
Stage 1
Stage 2
Stage 3
Selection of Items for Item Analysis is Made
Item Analysis Completed & Final Validity Study Begins

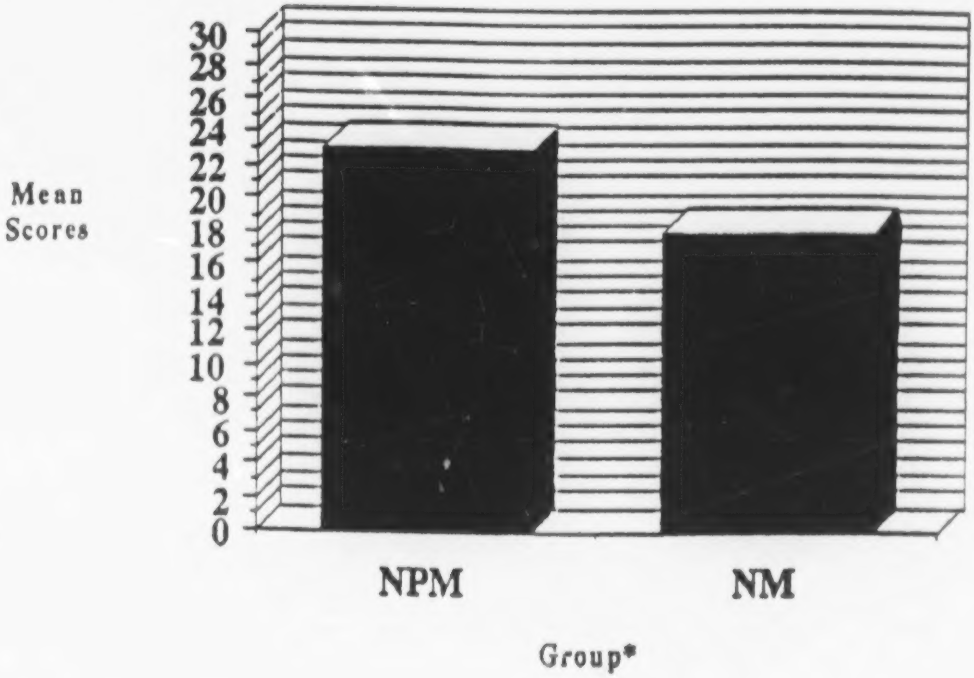
* Group or subgroup formed from professional master and nonmaster groups used in reliability and concurrent validity analysis.

Figure 1. Timeline showing three stages of the instrument's development.

Figure Caption

Figure Caption

Figure 2. Non-professional master and nonmaster mean scores in item analysis.

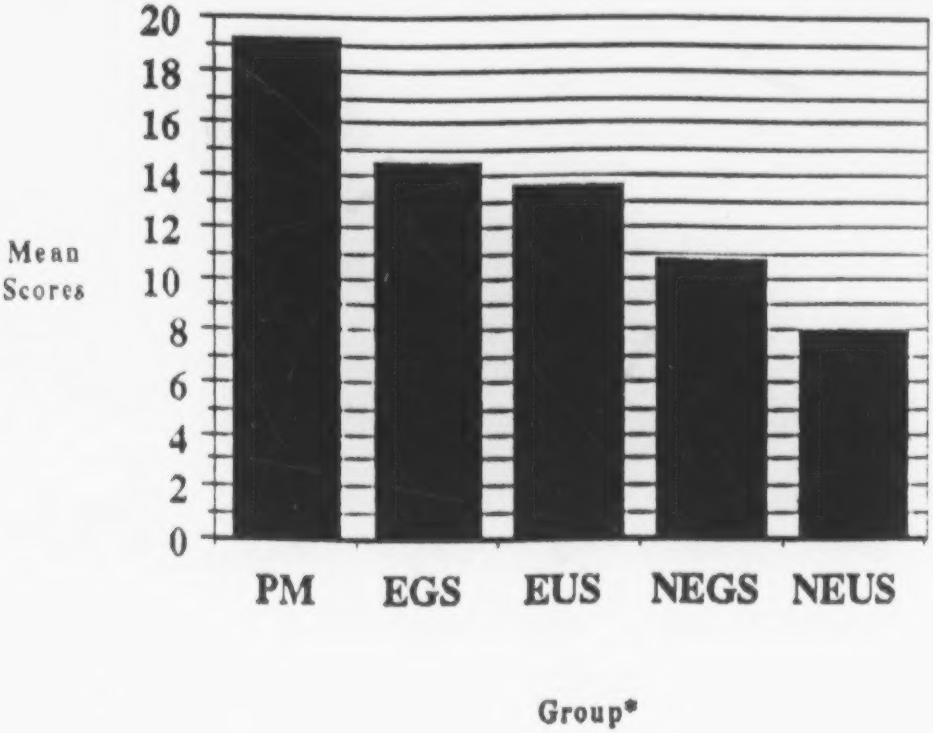


*NPM = Non-Professional Masters

NM = Nonmasters

Figure Caption

Figure 3. Professional master group and nonmaster subgroup mean scores.



*PM = Professional Masters

EGS = Education Graduates

EUS = Education Undergraduates

NEGS = Non-Education Graduates

NEUS = Non-Education Undergraduates

Figure Caption

Figure 4. Smoothed frequency distributions of masters and nonmasters.

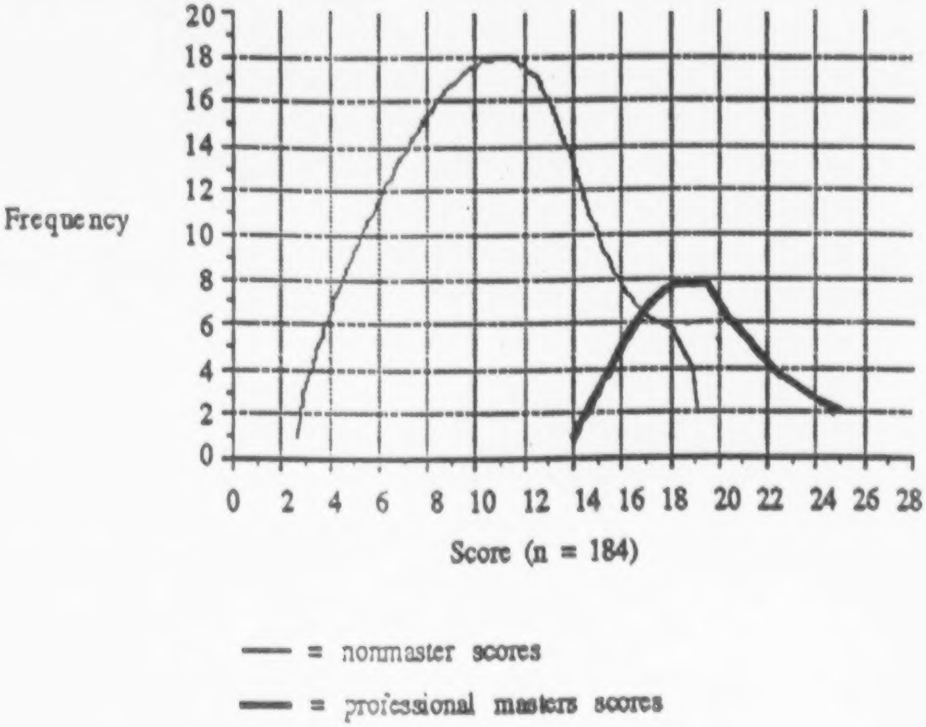


Figure Caption

Figure 5. Phi matrix at a cut off score of 17.

		Classification of Instrument		
		Nonmaster	Master	
Classification of Subject Matter Experts	Master	4 11.76%	30 88.24%	4 + 30 = 34
	Nonmaster	134 89.33%	16 10.67%	134 + 16 = 150
		4 + 134 = 138	30 + 16 = 46	Σ = 184

n = 184

Figure Caption

Figure 6. Actual and predicted frequencies produced by a discriminant analysis.

Actual Classification

	1	2	3	4	5	Total
1	34	0	0	0	0	34
2	2	30	6	5	0	43
3	0	2	14	2	5	23
4	0	8	1	35	1	45
5	0	0	7	4	28	39
Total	36	40	28	46	34	184

Predicted Classification

Figure Caption

Figure 7. Professional master and nonmaster subgroup mean scores (24 items).

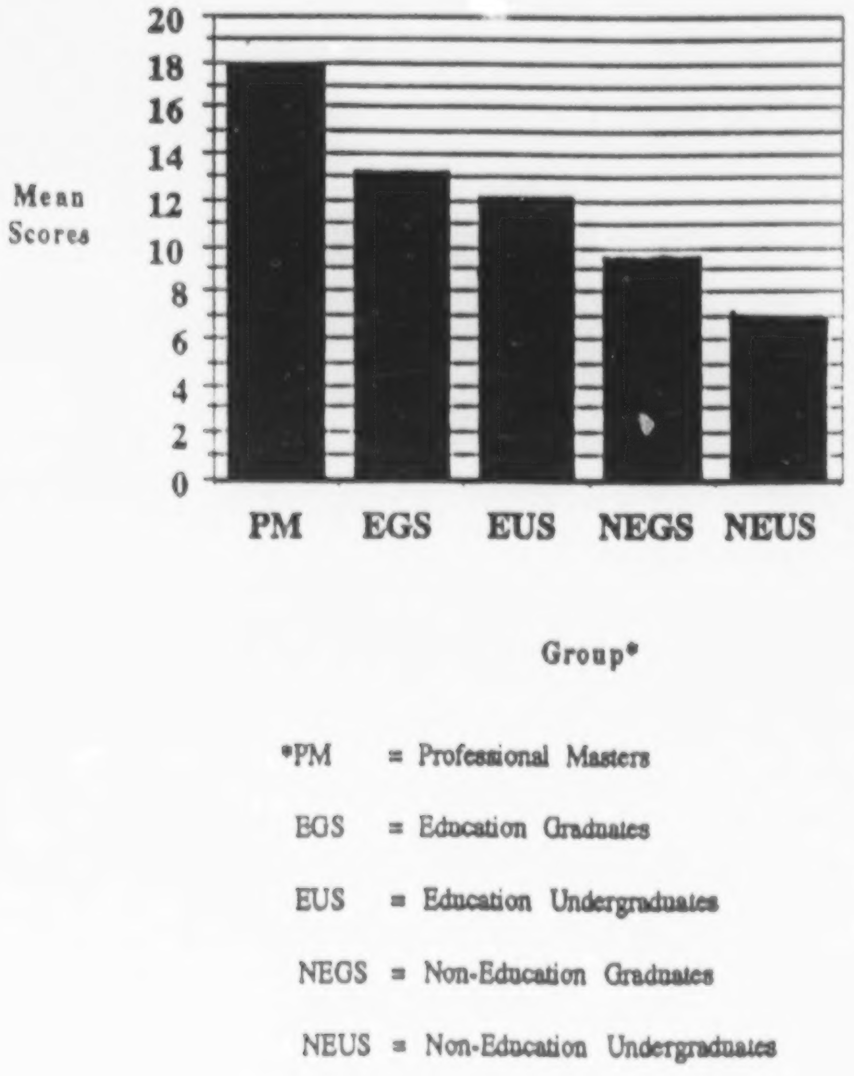
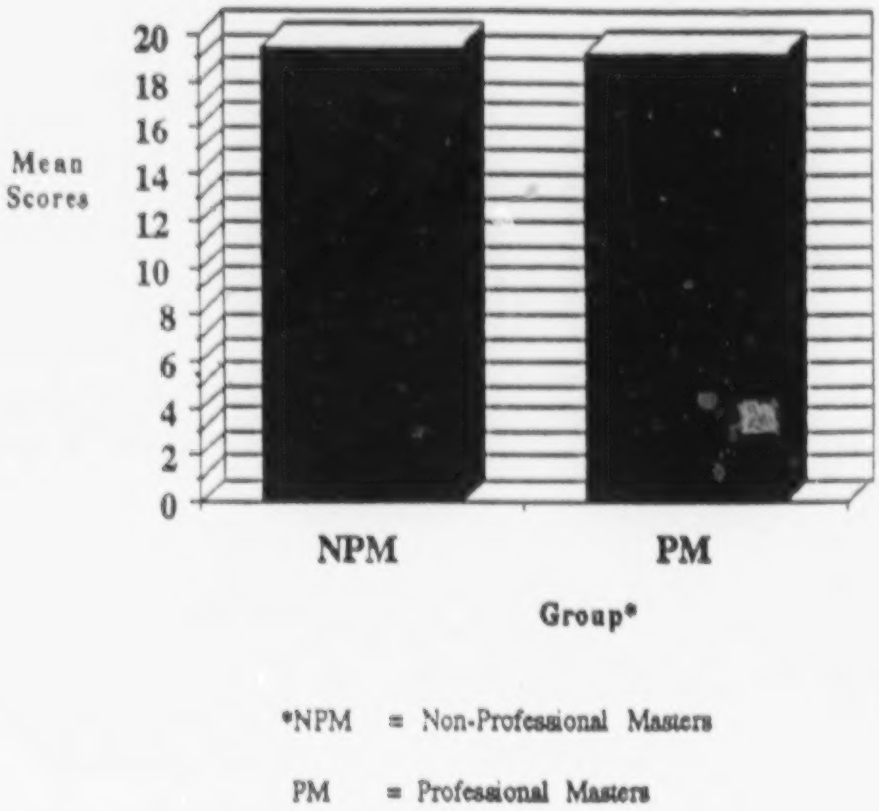


Figure Caption

Figure 8. Non-professional master and professional master mean scores (28 items).



Related References

Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole Publishing Company.

Boothe, B. (1984). Certification--beyond reason. *Performance & Instruction Journal*, *23*(1), 19-20.

Bratton, B. (1984). Professional certification: Will it become a reality? *Performance & Instruction Journal*, *23*(1), 4-7.

Briggs, L.J. (Ed.). (1976). *Instructional design: Principles and applications*. Englewood Cliffs, NJ: Educational Technology Publications, Inc.

Bureau of Labor Statistics. (1986). *Occupational projections and training data: A statistical and research supplement to the 1986-87 occupational outlook handbook* (Bulletin 2251). Washington, DC: U.S. Government Printing Office.

Coscarelli, W. (1984). Arguments for certification: A political "trip-lett? *Performance & Instruction Journal*, *23*(1), 21-22.

Davis, R.H., Alexander, L.T., & Yelon, S.L. (1974). *Learning system design: An approach to the improvement of instruction*. New York: McGraw-Hill Book Company.

Deden-Parker, A. (1979). Instructional design competencies for business and industry designer-client interactions. *Educational Technology*, *19*(5), 44-46.

Deden-Parker, A. (1981). Instructional technology skills sought by industry. *Performance & Instruction Journal*, *20*(1), 24-25, 30.

Feltovich, P.J., Johnson, P.E., Moller, J.H., & Swanson, D.B. (1984). LCS: The role and development of medical knowledge in diagnostic expertise. In W.J. Clancy, & F. Shortlesse (Eds.). *Readings in Medical Artificial Intelligence: The past decade* (pp. 275-319). Reading, MA: Addison-Wesley.

Fleming, M., & Levie, W.H. (1978). *Instructional message design: Principles from the behavioral sciences* (4th ed.). Englewood Cliffs, NJ: Inc.

Gagne, R.M. (Ed.). (1987). *Instructional technology: Foundations*. Hillside, NJ: Lawrence Erlbaum Associates, Publishers.

Galey, M. (1980). Certification means professional development. *Instructional Innovator*, *25*(9), 25-31.

Glass, G.V., & Hopkins, K.D. (1984). *Statistical methods in education and psychology* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

International Board of Standards for Training, Performance, and Instruction. (1986). *Instructional design competencies: The standards*. Iowa City: International Board of Standards for Training, Performance, and Instruction.

Laird, D. (1985). *Approaches to training and development*. New York: Addison-Wesley Publishing Company, Inc.

Logan, R.S. (1982). *Instructional systems development: An international view of theory and practice*. New York: Academic Press.

Mager, R.F. (1973). *Measuring instructional intent or got a match?* Belmont, CA: Fearon Publishers.

Markle, S.M. (1978). *Designs for instructional designers*. Champaign, IL: Stipes Publishing Company.

McCullough, M. (1981). ASTD professional development update. *Training and Development Journal*, *35*(7), 17-18.

Moses, J., & Byham, W. (1977). *Applying the assessment center method*. New York: Pergamon Press.

Norman, G. (1990). *Research in the psychology of clinical reasoning: Implications for assessment*. Manuscript submitted for publication.

Pinto, P.R., & Walker, V.W. (1978). What do training and development professionals really do? *Training and Development Journal*, *32*(7), 58-64.

Prigge, W.C. (1974, November). Accreditation and certification: A frame of reference. *Audiovisual Instruction*, 12-18.

Romiszowski, A.J. (1981). *Designing instructional systems: Decision making in course planning and curriculum design*. London: Nichols Publishing.

Schmidt, H.G., Norman, G.R., & Boshuizen, H.P.A. *A cognitive perspective on medical expertise: Theory and implication*. Manuscript submitted for publication.

Shrock, S.A., & Foshay, W.R. (1984). Measurement issues in certification. *Performance & Instruction Journal*, *23*(1), 23-27.

Spaid, O.A. (1986). *The consummate trainer: A practitioner's perspective*. Englewood Cliffs: Prentice-Hall.

Stolovitch, H.D. (1981). Preparing the industrial and educational instructional developer: Is there a difference? *Performance & Instruction Journal*, *20*(1), 29-30.

Swezy, R.W. (1981). *Individual performance assessment: An approach to criterion-referenced test development*. Englewood Cliffs: Prentice-Hall.

Tabachnick, B.G., & Fidell, L.S. (1983). *Using multivariate statistics*. New York: Harper & Row.

Task Force on ID Certification. (1981). Competencies for the instructional/training development professional. *Journal of Instructional Development*, *5*(1), 14-17.

Tiemann, P.W., & Markle, S.M. (1978). *Analyzing instructional content: A guide to instruction and evaluation*. Champaign, IL: Stipes Publishing Company.

Trimby, M.J. (1982). *Entry-level competencies for instructional developers*. Dallas, TX: Paper presented at the Annual Meeting of the Association for Educational Communications and Technology. (ERIC Document Reproduction Service No. ED 222 174)

Wallington, D.J. (1981). Generic skills of an instructional developer. *Journal of Instructional Development*, *4*(3), 28-32.

Westgaard, O. (1983). Certification for performance and instruction professionals: The solution lies in the use of assessment centers. *Performance & Instruction Journal*, *22*(7), 3-7.

Westgaard, O. (1984). Certification in instructional design really is an issue. *Performance & Instruction Journal*, *23*(1), 3.

Appendix A
References Used to Identify Instructional Design Competencies

Analyzing instructional content: A guide to instruction and evaluation	Tiemann, P.W. & Markle, S.M.
Designing instructional systems: Decision making in course planning and curriculum design	Romiszowski, A.L.
Designs for instructional designers	Markle, S.M.
Individual performance assessment: An approach to criterion-referenced test development	Swezy, R.W.
Instructional design: Principles and applications	Briggs, L.J.(Ed.)
Instructional message design: Principles from the behavioral sciences	Fleming, M. & Levie, W.H.
Instructional systems development: An international view of theory and practice	Logan, R.S.
Instructional technology: Foundations	Gagne, R.M. (Ed.)
Learning system design: An approach to the improvement of instruction	Davis, R.H., Alexander, L.T., & Yelon, S.L.
Measuring instructional intent or got a match?	Mager, R.F.
The consummate trainer: A practitioner's perspective	Spaid, O.A.

Appendix B
Example Questions in the Instructional Design Assessment Instrument

1. All things being equal, which of the following concept definitions would trainees learn more easily?
   a. A desktop publishing program is a program for importing word processing files and for page layout and for mixing graphics and text.
   b. A microcomputer system error occurs from sloppy programming or there is a memory error or there is a hardware error.
   c. Refined oil is thicker than water but not quite as thick as crude oil.
   d. An incorrect formative evaluation is an evaluation that was not done or it was done incorrectly or it was only partially done.

2. You have to develop training for a group of secretaries who will have to use a new word processing program to be used by the company. The company has not used computers for word processing before this time and a survey has shown that most of the secretarial staff have never used a computer. What would be the most economical and efficient instructional sequence for each part of the instruction?
   a. statement of a step, example of the step, another example of the step requiring a secretary response
   b. example of a step, statement of a step, another example requiring a secretary response
   c. statement of a step, example of the step requiring a secretary response
   d. statement of a step, restatement of the step requiring a secretary response

3. A trainer is concerned that he is talking too quickly for the learners to understand and take good notes. What would be the most appropriate method of collecting data to see if the trainer's concerns are valid?
   a. an open-note test
   b. an audio recording of the trainer's lecture
   c. a final course evaluation completed by the learners
   d. observation of the training sessions by another trainer

4. The final evaluation for a required training unit to teach telephone operators to use a new long distance dialing system consists of a questionnaire. Below are 4 questions from the questionnaire. Which of the 4 questions should be eliminated?
   a. Was the training helpful?
   b. Did the instructor ask a lot of questions?
   c. Were the computer simulations useful in learning the task?
   d. Were the workbook exercises useful?

Title:

Research on Hypermedia Browsers and Suggestions for Future Research

Authors:

William B. Story, Jr.
Francis A. Harvey

Research on Hypermedia Browsers and Suggestions for Future Research

William B. Story, Jr.
Francis A. Harvey
Lehigh University

Introduction

Browsing tools have been developed to help users make effective use of the large amounts of information in hypermedia applications, aiding the user both in orientation (where am I in the collection of information contained in the hypermedia application?) and in navigation (how do I get to another specific piece of information?). There are few formal research studies on how users actually utilize browsers in hypermedia. This paper describes current research on browsing tools and the types of browsers currently in use; describes why they are needed, and suggests additional research which might lead to the design of more effective browsing tools.

Browsing Defined

Browsing is the intellectual process of acquiring individualistic knowledge. Users browse through a knowledge base to find information that is individually meaningful (Jonassen, 1989). McAleese states that when browsing users are following a path or strategy until some goal is reached. (1989, p 11) The user's ability to browse quickly and easily through a hypermedia system is critical, especially for large-scale hypermedia applications, where it is also necessary to navigate through a large number of information nodes (Akscyn, McCracken, & Yoder, 1988). McAleese (1989, p. 7) makes note of a useful and necessary distinction between the terms browsing and navigating. He states:

> There are two states: browsing is where an idea is followed using the linking mechanism of the hypertext elements (e.g. card, windows, nodes); navigation involves the use of a graphic aid such as a browser or map to show an overview representation of the nodes and links.

A Taxonomy of Browsing Tools

Browsing tools (browsers) are utility functions built into hypermedia applications. We classify browsers according to two attributes: 1) mode (text or iconic); and 2) whether the form of the browser display represents the structure of the knowledge base contained in the hypermedia application.

Tables of contents or alphabetical indices are examples of text browsers. A table of contents is a structural text browser, since the sequence of topics in the table of contents represents one meaningful organization of the hypermedia material . An alphabetical index is a non-structural text browser, since the alphabetical listing of topics gives no indication of the underlying organization of the knowledge base.

The hypermedia knowledge base *In the Holy Land* (Optical Data Corporation, 1989) uses a non-structural iconic browsing tool. The row of icons displayed across the bottom of each screen, representing various topics in the knowledge base or utility functions (e.g. note-taking or videodisc controller), gives no indication of the organizational structure of the knowledge base.

The Cities in Schools hypermedia training materials developed at Lehigh University used a variety of browsing tools (Harvey and Story, 1990). A timeline representing the major phases of the development of Cities in Schools was used as a structural iconic browser for the history module. All modules used a non-structural hybrid text/iconic browser, where major sections and utility functions were represented by a set of text buttons clustered across the bottom of the screen.

Review of the Literature on Browsing and Browsers

A comprehensive review of the instructional design, educational technology research, and library and information science literature is now in progress, and will be reported in a later paper. Preliminary review indicates little agreement on terminology used to describe browsing tools. Such agreement is sorely needed.

Types of Browsers

The type of browsing tool most commonly found in hypermedia applications is an iconic browser, although in many cases these are non-structural. Structural iconic browsers are frequently referred to in the literature and are usually called maps. Fiderio (1988, p. 240) uses the term graphical browser to describe structural iconic browsers:

> The *graphical browser* is a node that contains a structural diagram of a network of nodes. While not all systems provide a *graphical browser*, most attempt to provide some type of overview system that helps you stay oriented in the network and visualize how information is linked. In large hypertext environments consisting of hundred of nodes, browsing tools are especially important because it's so easy to get lost.

Halasz (1988) describes graphical browsers (without specifying whether text or iconic representation is used) as usually containing a structural diagram of a network by which users can navigate and identify their location.

Structural browsers are often referred to as "maps" which represent "terrain" or "street" views of what has to be learned (McAleese, 1988). Maps and tables of contents aid the user in navigation by showing the overall structure of the knowledge base (Locatis, Letourneau, & Banvard, 1990).

Why Browsers are Needed

Kearsley points out the need to minimize the burden on the user's short-term memory (1988, p. 21): "The user should not be required to remember things from one screen to another." Fiderio discusses the problem and the proposed solution (1988, p. 239): "...You can forget how or why you got to where you are. To alleviate this problem of disorientation, many systems provide a tool called a ... *browser.* " Fischer, Mandl (1990, p xxv) point out:

The user should always be able to answer the following questions:

1. Where am I?
2. How did I get here?
3. What can I do here?
4. Where can I go to?
5. How do I get there?

One of the main problems for users of hypermedia is the sense of "getting lost." Dede describes this problem of cognitive overhead in these terms: "The richness of non-linear representation carries a risk of potential intellectual indigestion, lost of goal-directness, and cognitive entropy" (Dede, 1988, p. 8). Weyer points out that (1989, p. 98) "without help, browsing through everything in an information system can be as ineffective as searching blindly...." Structural browsers which display a diagram on which the user's location is highlighted provide instant orientation to the user, while at the same time allowing movement from one section to another within the application (Raker, 1989).

In a hypermedia application, there is usually too much information available for the user to comprehend. The user needs support to comprehend how much information is available and how it is interconnected. Browsing through vast information directly without any filtering device will be frustrating to the user (Weyer, 1989).

Suggestions for further study

The following are representative questions for formal research on the use of browsers:

- Which type of browser is most beneficial to specific types of users?
- Do novices require different browsing support from experts?
- For which applications are text browsers more appropriate than iconic browsers?
- For which purposes are structural browsers more appropriate than non-structural browsers?
- Are there interactions between the degree of structure included in a browser and the way information is represented (text vs. icons)?
- For very large or very complex hypermedia applications, what is the most effective strategy for "layering" successively more detailed browsing tools?

Hypermedia has great potential for knowledge access and utilization and for instruction. Careful and systematic research on specific aspects of hypermedia such as browsers and browsing will help make this potential a reality.

References

Akscyn, R. M.; McCracken, D. L.; and Yoder, E. A., KMS: A distributed hypermedia system for managing knowledge in organizations. Communications of the ACM, 31(7), pp. 820-834.

Dede, C. (1988). The role of hypertext in transforming information into knowledge. Paper presented at the annual meeting of NECC, Dallas, TX, June 1988

Fiderio, J. (1988). A grand vision. BYTE, 13(10), pp. 237-244.

Fischer, P. M., Mandl, H., (1990). Introduction: Toward a Psychophysics of Hypermedia. In D. Jonassen & H. Mandl (Eds.), Designing Hypermedia for Learning (pp XIX-XXV). New York: Springer-Verlag.

Halasz, F. G. Reflections on notecards: Seven issues for the next generation of hypermedia systems. Communications of the ACM, 31(7), pp. 836-851

Harvey, F.A. & Story, W.B. (1990). Designing, developing, implementing, and evaluating hypermedia training materials for human services professionals. Paper presented at the annual meeting of the Association for Educational Communications and Technology, Anaheim, CA.

Jonassen, D. H. (1989) Hypertext/hypermedia. Englewood Cliffs, NJ: Educational Technology Publications.

Kearsley, G. (1988, November). Authoring considerations for hypertext. Educational Technology, pp. 21-24.

Locatis, G.; Letourneau, G.; Banvard, R. (1990). Hypermedia and instruction. ETR&D, 37(4), pp. 65-77

McAleese, R. (1988). From concept maps to computer based learning: the experience of notecards. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

McAleese, R., (1989). Navigation and browsing in Hypertext. In R. McAleese (Ed), Hypertext theory into practice (pp 6-44). Norwood, New Jersey: Ablex Publishing Corporation

Optical Data Corporation (1989). *In the Holy Land*. Warren, NJ: Optical Data Corporation.

Raker, E. J. (1989). Hypermedia: New technology tool for educators. The Computing Teacher, 17(1), pp. 18-19

Weyer, S. A. (1989). As we may learn. In S. Ambron &.K. Hooper (Eds.). (1988). Interactive multimedia: Visions of multimedia for developers, educators, & information providers. (pp. 89-103). Redmond, WA: Microsoft Press.

Title:

Instructional Design and Human Practice: What can we Learn from Habermas' Theory of Technical and Practical Human Interests?

Author:

Michael J. Streibel

Instructional Design and Human Practice:
What Can We Learn From
Habermas' Theory of Technical and Practical Human Interests?

INTRODUCTION

This paper is going to answer a question that was first asked more generally about curriculum by Shirley Grundy: Is instruction a product or a practice (Grundy, 1987)? The answer to this question will have profound implications for instructional designers because if the outcome of instructional design is a product (e.g., an instructional system or a courseware program) then the work of instructional design is done before instruction begins. If, on the other hand, instruction is a form of practice, then the work of instructional design is integral to the practice of instruction.

To help answer our question about product or practice, I will base many of my arguments on those made by Grundy about curriculum. I will also build on the ideas of Koetting, Nunan, Bullough, and Nichols (Koetting, 1979; Nunan, 1983; Bullough et al., 1984; Nichols, 1989). Of course, ultimately, these ideas trace back to Habermas' Theory of Human Interests (Habermas, 1972, 1984, 1987). Habermas' framework will be explained throughout the paper. However, for now, it is important to clarify the assumptions that Habermas makes about the relationship between theory and practice.

The first assumption is something that every practicing teacher knows: Instructional theories do not determine their instructional practice in any significant way. Some researchers like Heinich (1988) lament that fact. Others like myself find this fact to be a fundamental truth about the human condition (Streibel, 1986, 1988). Namely, all human practice (whether it be the work of instructional designers, or teachers, or human learners) is situated in an ongoing context that requires continual judgment (Streibel, 1989). Hence, instructional theories are usually treated as resources rather than plans by practitioners and are evaluated for authenticity rather than productive value by practitioners. Grundy summarizes this position very well when she states that "theoretical explanations (e.g., theories of instruction) ... [have to be] grounded in the reality of teachers' experiences." (Grundy, 1987, p. 3). The same is true for the relationship of theories of instructional design and the lived reality of instructional designers, and, for the relationship of theories of learning and the lived reality of human learners (Streibel, 1989).

The second assumption that needs to be clarified is that grounding theories in the lived experiences of persons is not the same thing as:

1. inferring theories from experience (that would assume that the theoretical components of experience are in the experience and only waiting to be discovered).
2. mapping theories onto experience (that would treat theories as totally artificial "useful fictions").

Rather, experience contains within it various transcendental realities that have the potential for being articulated with socially-constructed symbols and being realized in social practice. I will say more about this phenomenon later in the paper.

You can already see from what I have said that I have begun to address the question of whether instruction is a product or a practice. Grundy provides a useful insight on this point (Grundy, 1987, p.5):

> Curriculum [she writes] is not a concept, it is a cultural construction ... a way of organizing a set of human educational practices.

The same can be said for instruction. Instruction is not a concept but a "way of organizing a set of human practices." Otherwise, we would have a technical approach to education because:

1. the concept comes first (eidos - e.g., objective).
2. the instructional plan comes second.

3. the implementation of the plan comes third.
4. the product comes fourth.
5. the evaluation of the product comes last.

Figure 1 About Here

Where are the people in this picture? Everything seems to be about pre-existing ideas which become explicit plans that then control specific actions. Teacher's behaviors and learner's behaviors are always compared against pre-existing plans and objectives. If, on the other hand, instruction is conceptualized as a set of cultural practices, then we can deal with how people experience the emergence of an intention and how they take an active and responsible role in bringing about a goal. Remember, when I am talking about people here, I am referring as much to instructional designers as I am about teachers and learners when each of these persons do their respective work.

Figure 2 About Here

A further clarification is needed about the idea of instruction as a cultural construction of practice. To claim, as Grundy does, that the curriculum (or instruction in our case) is a social construction is not the same thing as claiming that there is a social component to the curriculum (or instruction). Otherwise, social factors would be just another set of factors that are external to the instructional practices of people.

Figure 3a&b About Here

Grundy's clarification about this issue for curriculum applies equally well to instruction (Grundy, 1987, p. 6):

> To think about curriculum [she says] is to think about how a group of people act and interact in certain situations. It is not to describe and analyze an element which exists apart from human interaction.

This is a call for a different type of theory and a different type of relationship between theory and practice. It is a call for an elevation of practice to an equal status with, and a dialectical relationship with, theory.

Jurgen Habermas provides the most complete articulation of a framework that deals with different types of relationships between theory and practice. I will use his ideas as well as Grundy's application of these ideas to curriculum as guides for dealing with instructional design, instruction, and learning.

According to Habermas, there are three types of human interests which ultimately manifest themselves as three ways of knowing. Grundy describes these three human interests as follows (Grundy, 1987. pp. 6-19):

1. The technical human interest entails empirical/analytical ways of knowing that represents the world in terms of objects, processes, and laws which describe the transformation of objects and processes. The natural sciences, for example, display a fundamental technical human interest. In some cases, instructional theories which are constructed and applied as if they were a natural science also fall under the umbrella of the technical human interest (Reigeluth, 1983, 1987; Gagne, 1987; Gagne et al. 1988).
2. The practical human interest entails historical and hermeneutic ways of knowing that represent the physical, social, and cultural worlds as "texts" which have to be interpreted in order for meaning to emerge. Meanings, however, are not in the texts just waiting to be discovered and decoded. That would revert to the technical approach. Rather, meanings are socially and culturally constructed in the speech acts and practices of interacting human agents.

3. Finally, the emancipatory human interest entails a critical way of knowing where critical theorems are gleaned through collective reflection on social and cultural practices and then used to restructure future actions. The struggle here is to:
    a. become conscious of the pre-understandings in existing social and cultural practices,
    b. uncover the contradictions between the ideals of truth, justice, and freedom and actual social and cultural practices,
    c. change social practices.

Figure 4 About Here

Although I will describe each of these three human interests in greater detail below, I will only apply the technical and the practical human interests to instructional design in the rest of the paper.

The Technical Human Interest

The technical human interest, according to Habermas, entails one way of relating to the world. The best example of this comes from the empirical-analytical sciences where a neutral observer (e.g., a researcher) makes "positive" (i.e., publicly verifiable) observations and then posits hypothetical relations about the regularities in the objects and processes of the world. This eventually results in lawful statements about the world that are extended to ever more abstract levels on the one hand and ever more detailed observations on the other. The lawful statements then permit prediction of future observations. This opens up the possibility for linking predictions of future observations with control of future observations because descriptive theories lend themselves to prescriptive use. Reigeluth in his book on Instructional Theories makes this very point (Reigeluth, 1983). The problem with shifting from description to prescription, however, is the semantic shift that takes place. Habermas, quoted in Grundy, makes this point perfectly clear (Grundy, 1987, p. 11): "The meaning of such predictions [writes Habermas] ... [becomes] their technical exploitability." Prediction becomes control of future actions and control becomes part of the semantic content of knowledge. This, in turn, leads to instrumental action which is guided by technical rules. Grundy summarizes these ideas as follows (Grundy, 1987, p. 12):

> The technical interest is: a fundamental interest in controlling the environment through rule-following action based on empirically grounded laws.

These consequences are acceptable if we are dealing with phenomena that lend themselves to technical exploitability. But what about the case where we use instructional theories to prescribe future instructional actions for sentient human teachers, or we use learning theories to prescribe future learning actions of sentient human learners? The prescriptive use of instructional or learning theories may violate the very way that teachers and learners act in the world (unless they are willing to act like autonomatons, or more generally, like information processors) (Streibel, 1986). This is not to suggest that empirically-grounded instructional or learning theories are totally useless knowledge. They are useful knowledge to teachers (and learners and instructional designers) if they are used as resources rather than controllers for future actions. An Instructional Designer therefore has to find ways to design resources rather than plans for teachers and learners.

The Practical Human Interest

The practical human interest, according to Habermas, entails a different orientation to the world than the technical human interest. Whereas the technical interest is bent on controlling the world, the practical interest is focused on understanding the world. The reason is simple: human beings not only want to be in the world (rather than neutrally observing, theorizing about, and controlling the world) but they have to be in the world. It is part of our human condition. Understanding here, according to Grundy, is more than technical understanding, but the creation of meaning in our lives. Practical understanding therefore has to come to terms with our lived experience in the world. This form of understanding is best developed by the historical-hermeneutical sciences. Note that in this approach to the world, facts are no longer out there to be discovered but are accessed through our understanding. Polanyi made a similar point when he claimed that we can only focus on explicit facts and knowledge through our tacit knowledge (Polanyi, 1958).

This brings up an interesting dilemma. What exactly does a community of persons examine or look at in the process of socially constructing meanings and knowledge? One person cannot "get into" the mind and heart of another and therefore does not have direct access to another's experience. Furthermore, every individual's experience cannot be the bases for the reconstruction of all knowledge by each individual. The very fact that a community of interacting persons exists argues against a Leibnizian form of individualism. Grundy resolves this dilemma by saying that "both empirical and interpretive sciences have to transform human action into something else in order to study it" (Grundy, 1987, p. 13). In the case of the empirical sciences (of which instructional science is part), human actions are converted into "behaviors." In the case of the hermeneutic sciences, human actions are converted into "texts."

Figure 5 About Here

Note, however, that the "text" of human action is not treated as if it had meaning in it. Hermeneutics is not a form of technical rationality (i.e., it is not a matter of applying procedures). Hermeneutics is a matter of situational judgments.

What counts as success in the realm of practical action? Grundy again spells out the criteria for success (Grundy, 1987, p. 14):

> Knowledge which is concerned with understanding is not to be judged according to the success of the operations arising as a consequence of that knowledge. Rather, it is to be judged according to whether the interpreted meaning assisted the process of making judgments about how to act rationally and morally [in the world].

Notice the juxtaposition of technical and practical action. In technical action, the consequence of action is a product (a behavioral outcome in the case of an instructional action) that is compared against a pre-existing idea (e.g., an objective). In practical action, the participants in the action themselves (e.g., the teachers or learners) have to make an on-going series of judgement calls about whether they are moving towards greater understanding. This never-ending goal contains both a technical and a moral dimension (i.e., whether the learner's future action matches their original intention and whether the learner's future action is worthwhile and good for them -- something that cannot be predicted). The knowledge that a teacher and a learner needs is worked out during a series of judgement calls where interpreted meanings are authenticated by the teacher and the learner in the actual unfolding situation.

The best instructional example I ever heard of this was where a teacher showed a video tape about alcoholism in a health class. The teacher proceeded on the assumption that the students would use the examples in the tape as evidence for the evils of alcohol. They did! However, many of the students also sympathized with the alcoholics in the video because they realized that the alcoholics were in an impoverished economic environment where drinking was one of the few avenues left to maintain a modicum of control. If the teacher had not caught on to the meanings that the students were bringing to this situation, the intended health lesson and subsequent evaluation of what was learned from the tape would have been a totally lost. The point of this example is that an instructional designer who spells out a learning objective cannot predict what meanings potential learners will bring to the learning situation. Hence, an instructional designer cannot rely on a technical approach to design. Rather, an instructional designer has to be guided by a practical human interest and support the instructional and learning processes that actually take place. Grundy again spells out the implication of the practical interest (Grundy, 1987, p. 14):

> The practical interest [she writes] is a fundamental interest in understanding the environment through interaction based on a consensual interpretation of meaning.

The Emancipatory Human Interest

Habermas' third human interest goes beyond the technical and the practical by focusing on the ways that people struggle to change their social, economic, and cultural conditions of existence towards forms that are more truthful, more just, and more free. Habermas claims that the ideas of truth, justice and freedom are transcendental realities within everyday human interaction rather than within some external reality. Emancipation, therefore, is not an abstract, external idea but a potential waiting to be realized in the world of human beings. However, to realize emancipation, human beings have to:

1. become critically conscious and aware of how they construct their current knowledges, beliefs, and practices.
2. socially-reconstruct their knowledges, beliefs, and practices.

The technical way of framing knowledge, beliefs, and practices will not help here. Grundy again describes the reasons (Grundy, 1987, p. 17):

> The technical interest will not facilitate autonomy and responsibility because it is an interest in control. An interest in control will certainly facilitate independence for some, but this is a false autonomy, for it is an 'autonomy' which entails regarding fellow humans and/or the environment as objects.

The technical approach to instructional design will therefore not lead to those types of knowledges, beliefs and practices that enhance the emancipatory potential of human beings no matter how efficiently or effectively it produces learning outcomes. Emancipation, or empowerment as Grundy claims it is now called in educational research literature, is "the ability of individuals and groups to take control of their own lives in autonomous and responsible ways" (Grundy, 1987, p. 19).

What will an emancipatory curriculum or emancipatory instruction achieve? According to Grundy, at the level of consciousness (Grundy, 1987, p. 19):

> subjects participating in the educational experience will come to know theoretically and in terms of their own existence when propositions represent distorted views of the world ... and when they represent invariant regularities of existence.

At the level of practice (Grundy, 1987, p. 19):

> the educational encounter ... [will include] action which attempts to change the structures within which learning occurs.... [This will] entail a reciprocal relationship between self-reflection and action.

The emancipatory human interest can therefore be defined as (Grundy, 1987, p. 19):

> a fundamental interest in emancipation and empowerment to engage in autonomous action arising out of authentic insight into the social construction of human society.

I will not address the effect of an emancipatory human interest on instructional design in this paper but will leave that for a future study. Other researchers have already begun to reconstruct instructional design in light of emancipatory human interests (Koetting, 1979; Nunan, 1983; Nichols, 1989) as have educators begun to reconstruct pedagogical practice into critical pedagogical practice (Freire, 1970; Shor, 1980; Livingston, 1987). I therefore now turn to a discussion of treating instruction as a product of instructional design (the technical human interest approach to instructional design) and treating instructional design, instruction, and learning as various forms of inter-related practice (the practical human interest approach).

INSTRUCTION AS A PRODUCT OF INSTRUCTIONAL DESIGN

I would now like to apply Grundy's interpretation of Habermas' ideas to instructional design and instruction. Grundy's summarizes the relationship between theory and practice within the three human interests. This provides a good starting point for our discussion. I will then look at the implications of treating instruction as a product of instructional design and treating behaviors as the product of instruction.

The technical human interest approach to theory and practice treats theory as a guide to action. Traditional instructional design falls into this category. For example, when instructional designers use Reigeluth's Elaboration Theory or Merrill's Component Display Theory in a prescriptive manner, they are taking a technical approach to instructional design and treating instruction as a product of design (Reigeluth, 1983).

The practical human interest approach to instructional design, on the other hand, seeks to authenticate theories "in a process of self-reflection through with [one] test[s] the theoretical explanations in light of ... [one's] experience" (Grundy, 1987, p. 21). Hence, instructional designers in this latter view do not so much create a product called instruction as much as create resources for instructional practitioners in the instructional situation. These resources are then used by instructors as well as learners in a way that will be described in greater detail later in this paper.

The technical human interest approach to solving instructional problems follows a general sequential pattern as described in Figure 1. Notice that the technical approach to designing instruction takes for granted that an outcome an be predetermined and that a set of design plans and implementation plans can guide design work and implementation work to create the desired outcomes. The outcomes are then compared against the pre-existing objectives to see if they measure up. This whole approach turns the design process into a rational decision-making process where options and alternative paths are defined within a problem space (Streibel, 1989). Likewise, teaching becomes an instructional management problem of steering the learner through an imaginary instructional problem space and keeping records of his or her "progress." Instruction here is not a process of negotiating the meaning of means and ends with learners. Instructional design, instruction, and learning, in fact, are merely matters of skill.

Figure 6 About Here

The technical human interest approach to instructional design (and instruction) has some fundamental problems. First of all, this approach overlooks the fact that instructional designers are constantly making judgement calls in the process of crafting the instructional materials or the instructional systems. It also overlooks the fact that teachers (or students) who use these materials are constantly bringing their own meanings to the teaching (or learning) situation. However, these judgement calls on the part of designers (or teachers or students) are not conceptualized or legitimized within the technical approach to instruction. When they are acknowledged, they are given secondary or peripheral status such as subjective or social factors. Grundy describes this problem as follows (Grundy, 1987, p. 23):

> Although skilled actions may allow for some decision-making and choice, the range of choice, and hence the freedom that the artisan has to take action, is always restricted by the eidos [i.e., objective] of what is to be created.

The work of instructional designers (and of instructors) within the technical approach is therefore essentially reproductive. That is, designers are (Grundy,1987, p. 26):

> reproducing in the material world eideis which already exist in the abstract world of ideas or which have clearly been reproduced elsewhere.

Grundy goes on to describe the hierarchical relationships between theory and practice (Grundy, 1987, p. 27):

> The technical interest [she concludes] presupposes a hierarchical relationship between theory and practice. Practices exist in order to bring certain plans to fulfillment. Moreover, good practice is taken to be evidence of sound theory.

Hence, in the case of instructional designers who use a technical approach to instruction, instructional design practices exist in order to bring a model of instructional design to fulfillment. In the case of teachers, instructional practices exist in order to bring a model of instruction to fulfillment. In the case of the learner, instructional stimuli and learning activities exist in order to bring a model of learning to fulfillment. In each of these cases, the outcome is seen as a product and the model is believed to have a causal relationship to the product. The people (designers, teachers, and learners) are only implementors of the models.

What's wrong with the technical approach to instructional design and instruction? After all, this approach predominates our current thinking about design and instruction. Why change something we have been doing for decades? Why not continue to let theoretical statements "stand in a deterministic relationship to the world of practice" -- especially since theories are becoming much more refined? (Grundy, 1987, p. 28) Why not take advantage of this approach with computer-based instruction where theoretical models control the "behavior" of computer systems (Streibel, 1988). Grundy again provides a simple answer. In order for the technical approach to work "we must control both the learning environment and the learner" (Grundy, 1987, p. 29). In the case of control of physical objects, this does not cause a problem because it does not change their nature. However, in the case human beings, a controlling orientation changes their nature. Humans, after all, grow up into our images of the "other." What is, therefore, ultimately wrong with the technical approach to education, is that it embodies certain "power relationships within the learning environment" -- power relationships that do not lend themselves to the growth of autonomy and responsibility in the identity of the learners.

The technical orientation towards the design of instruction has another potential limitation. Because curriculum-making power is no longer vested in the teacher, the teacher only has implementation choices left. This constitutes a deskilling of teachers (Apple, 1975, 1979, 1982). Furthermore, the students also have no curriculum-making power. Finally, even the instructional designer is disempowered because (Grundy, 1987, p. 32):

> once the design process is complete, the plan becomes external to the planner, and has an authority which is separate from the person of the designer.

Each of these people in their respective spheres are deskilled in practice-making activities and reskilled in technological activities (e.g., designer becomes design-process manager, teacher becomes instructional manager, and learner becomes time-on-task manager). None of these people, therefore, "remain immune to the technologizing and obsolescence at work in the technological society." (Grundy, 1987, pp. 33-34).

Finally, there are problems with how the curriculum itself is shaped by the technological orientation. Grundy is very explicit (Grundy, 1987, p. 34):

> The technical interest ... promote[s] a view of knowledge as sets of rules and procedures or unquestionable 'truths'. Knowledge is regarded as a commodity, a means to an end.

Furthermore, she concludes that (Grundy, 1987, p. 35):

> the technically informed curriculum is not only bound by the culture of positivism as far as the selection of content is concerned, but the methodology by which the content is imparted is also determined by positivistic requirements about objectivity and outcomes.

The technological orientation, therefore, contains a positivistic epistemology (i.e., knowledge is made up of facts, laws, and procedures), an objectivist ontology (i.e., the world is made up of interacting objects with objective behaviors, cognitive structures, and skills), and a positivistic methodology (i.e., following instructional plans brings about a learning product).

How does evaluation fit into this picture when technological interests predominate. Evaluation, as was mentioned earlier, entails measuring the product of instruction against a pre-existing idea (e.g., an objective). To do this, both the learner, the learning process, and the learning outcome have to be objectified. However, as Grundy concludes, this "trivializes the teaching-learning act" (Grundy, 1987, p. 37) because it views both teaching and learning as mechanistic acts and because it hides the political nature of evaluation (Guba & Lincoln, 1989). Have you ever asked yourself, for example, what grading on a curve contributes to learning? Why can't everyone get an A under these conditions. Curriculum scholars who study this issue have concluded that grading on a curve has more to do with management and control of students than with fulfilling the learning potential of each individual (Kliebard, 1987). Finally, lest you think that individualized instruction is the answer to grading on the curve, individualized instruction carries the objectification and social fragmentation of learning one step further and, therefore, is an intensification of the technological approach to instruction (Streibel, 1988).

So! Where are we with respect to the technical orientation to instruction? Grundy sums up the technical interest as follows (Grundy, 1987, p. 52):

> The practitioner [and this could be the instructional designer, the instructor, or the learner] whose knowledge is constituted by the technical interest perceives the external eidos [e.g., objective] as a finite plan, and uses his/her skills to modify, adapt, and apply it in a different situation to produce an outcome that is judged in terms of efficiency and effectiveness.

The actual instructional designers, or teachers, or learners engaging in their respective practices (i.e., designing, teaching, learning), however, (Grundy, 1987, p. 52):

> grasp the eidos in terms of principles, relying upon practical judgments as a basis for decision. What is important for him/her is understanding and the creation of a meaningful learning environment.

Hence, instructional designers craft a design (whether instantiated in print, in video, or on a computer) that teachers use as a resource rather than as a plan. Teachers, in turn, craft an appropriate learning environment for each new learner. The learner then engages this environment to construct meaningful knowledge and actions under the guidance rather than the control of the teacher (Rogoff & Lave, 1984; Lave, 1988; Brown, 1988; Brown et al., 1989). Each of these people may pretend that they are following plans but, in fact, these plans are only resources for action (Suchman, 1987). The actual work of an instructional designer, an instructor, or a learner is therefore informed by a practical interest because they construct the means as well as the ends in the process of doing their work.

INSTRUCTIONAL DESIGN AND INSTRUCTION AS PRACTICE

I will now move beyond the technical framework and look at curriculum and instruction from the perspective of Habermas' theory of practical human interests. The term practical here emphasizes the stiuatedness of all human actions be they a designer designing instruction, a teacher teaching a lesson, or a learner learning something. The key to success for practitioners in each of these situations is human judgement which in turn is dependent on "reading" the "meaning" of situations. This elevates the importance of hermeneutical interpretation in work and learning. "Hermeneutical knowledge," writes Grundy, "is a pre-eminent form of knowledge upon which action can proceed." (Grundy, 1987, p. 59)

Grundy traces Habermas' notion of practical action back to Aristotle's notion of phronesis. Phronesis entails practical judgement and situational knowledge. Phronesis also involves taste which (Grundy , 1987, p. 61):

> has to do with what's fitting for a particular situation.... [Furthermore] knowledge, judgement, and taste combine to produce a discernment that is more than a skill.

Hence, practical judgement goes beyond technical decision-making because it contains (Grundy, 1987, p. 62):

> a disposition [that is oriented] towards 'good' rather than 'correct' action [as is technical action]. It possesses an aspect of moral consciousness which the disposition of techne [i.e., the technical orientation] lacks.

These ideas apply directly to designing and teaching and learning. For example:

1. when we force an instructional designer to follow instruction design models or procedures in order to achieve a predetermined design goal (as in expert instructional design systems) we restrict the designer to technical forms of rationality and deny them the opportunity to exercise practical (design) judgement.
2. when we force an instructor to follow instructional models or procedures in order to achieve a predetermined instructional goal (as in prescriptive instructional design theories), we restrict the instructor to technical forms of rationality and deny them the opportunity to exercise practical (instructional) judgement.
3. when we force a learner to follow learning models and procedures in order to achieve a predetermined learning outcome (as in cognitive learning theories), we restrict the learner to technical forms of rationality and deny them the opportunity to exercise knowledge construction judgments.

The lived reality of designers, instructors, and learners, however, indicates that these persons include practical judgments in their respective spheres. Why not design instruction to acknowledge this basic fact and stop restricting these persons to technical forms of rationality? The reorientation involves becoming clear about how to design with the "good" of persons in mind rather than with the "correct" learning outcomes in mind. What does this mean for instructional designers? Grundy again points the way to an answer.

The first thing to clarify is the different dispositions involved in the technical and the practical orientations. In the technical human interest, skill is "product related" and "works towards an end other than itself" (Grundy, 1987, p. 62) (e.g., instructional design skill working towards a predetermined instructional skill). In Habermas' practical human interest, on the other hand, practical judgement is "directed towards the process of taking action." (Grundy, 1987, p. 62). Figure 7 summarizes these different type of dispositions.

Figure 7 About Here

This constant judgement making is guided by a qualitative idea of the "good". The "good", however, is not something external and prior to the situation. Rather, it is "always in a state of being formed" and implicit in the situation. Hence, Grundy concludes that (Grundy, 1987, p. 63):

> since what is right cannot be fully determined independently of the situation, practical action is characterized by choice and deliberation.

Instructional designers therefore have to be willing to trust their sense of the "good" as they craft something for teachers and learners. They also have to create something that leaves a space for teachers and learners to apply their own sense of the "good" in the teaching and learning situation. Designing "teacher-proof" instruction does not leave such space for teachers nor does designing "idiot-proof" instruction leave such space for learners.

A second thing to clarify is that practical action "should be taken on the basis of a thorough understanding of the situation." (Grundy, 1987, p. 65) This is not achieved by way of empirical-analytical knowledge alone because such knowledge is not situation specific. Rather, the participants in a situation have to engage each other and the situation as they reflect on and deliberate about further action. Deliberation, however, says Grundy (Grundy, 1987, p. 65):

> incorporates processes of interpretation and meaning making of a situation so that appropriate action can be decided upon and taken. Appropriate action ... further[s] the good of the participants in the action.

Appropriate action does not, as in the case of technical action, serve someone's interests external to the situation (e.g., where an objective serves an institutional need). Rather, appropriate action subsumes technical forms of action because it is the participants in a situation who decide what serves their own good.

A third thing we have to clarify is the role of understanding in the practical orientation. Grundy spells out the parameters of situational understanding (Grundy, 1987, p. 67):

> In trying to understand anything, we come to it with certain predisposition and fore-meanings (pre-judgments or prejudices). The process of understanding or interpreting text is a process of allowing our prejudices (pre-judgments) to interact with the meaning that the author of the text intended so that the text becomes 'meaningful.'

Remember that the notion of "text" here encompasses both human interaction in situations as well as the more traditional notion of text. Hence, coming to understand a situation requires engagement with the situation and dialogue with others about the situation. This, in turn, means that practical action entails the "negotiation" of meaning by the participants in a situation. Meanings are not out there in situations waiting to be discovered and decoded. That would treat meaning making as a technological enterprise. Rather, meanings are interpersonally constructed.

Finally, negotiation presupposes the "equality of participants." Otherwise, the power imbalance between participants would predetermine the outcome of meaning making. The implications of the foregoing ideas for instructional designers are enormous:

1. the power relationship between employer (e.g., institutions) and instructional designers has to be addressed if the practice of instructional design is to include negotiated meanings and deliberated actions.
2. the power relationship between instructional designers and instructors has to be addressed so that the instructor in the teaching situation and not the instructional designer (who is not in the learning situation at the time of learning) determines the constructed meanings in the situation.

3. the power relationship between instructors and learners has to be addressed so that the learner has some space to construct his or her own meanings and understandings.

Notice that this issue is relevant even when the instructional content is classical physics or mathematics (where professional communities have worked out the meanings and symbolic notations for those meanings). For the novice learner, an understanding of the meanings of classical science and mathematics is constructed and negotiated anew each time within the learning situation. Hence, in order to respect the meaning-making processes of novice learners, neither the "text" of the instructional design (e.g., the instructional system) nor the "text" of the teacher's instructional actions has the authority to impose a set of meanings. Grundy summarizes this point as follows (Grundy, 1987, p. 69):

> [Curriculum as practice] reject[s] as legitimate educational content that which does not have at its heart the making of meaning for the learner. It is not sufficient [she continues] that the teacher is able to interpret the curriculum (or instructional) texts to come to an understanding of what the document prescribes.

Rather, the teacher (as well as the instructional designer) is only one of the many agents who engages the learner in the construction of meanings. Having a teacher who can decode the social and institutional prescription of curriculum and instruction certainly helps but it is not sufficient to bring about meaningful learning in the learner. Engagement, dialogue, and negotiation with the learner in specific situations is necessary. Grundy therefore concludes that (Grundy, 1987, p. 70):

> it is no longer makes sense to speak of evaluating the effectiveness of curriculum (or instruction) in terms of pre-specified objectives.

Pre-specified objectives are external to the work of teaching and learning. The worthwhileness of curriculum and instruction in a given situation (i.e., meaningful learning for the learner interacting with the teacher) has to be worked out and negotiated between the participants in that situation. A teacher can certainly wear multiple hats (i.e., represent society's interests, represent institutional interests, represent personal or professional interests, etc.) However, a teacher has an obligation to make sure that none of these interests overwhelms the learner's growth in understanding.

What do these ideas imply for instructional designers? Is there anything they can do in their own realm of practice that does not dominate the constructed meanings of teachers and learners? Grundy, quoting Stenhouse's research, supplies an answer (Grundy, 1987, p. 80): [practical curriculum development] should focus on supporting teachers in the exercise of their judgments in their local contexts.Hence, instructional designers should:

1. focus on teachers's experiences,
2 support the processes that teachers undertake to act meaningfully in a given situation,
3. create resources or environments that support teachers in the use of their judgments to improve their practices.

Teachers, after all, are the professional practitioners on the spot in the learning situations who have to figure out what things mean to learners at any moment.

Grundy continues by claiming that teachers only learn from their experiences when they systematically reflect on their experience and develop more refined practices. For teachers to mature as professional practitioners, they have to construct their own personal ways of dealing with increasingly complex situations. However, the plans and procedures they do construct do not have any authority over other teachers who might choose to follow similar plans (Suchman, 1987). Otherwise, we are back to the technological approach to teaching where expertise resides in the procedures and not in the personally constructed practices of the teacher (Dreyfus & Dreyfus, 1986; Streibel, 1989).

CONCLUSION

The main conclusion of this paper is that Habermas' theory of practical human interests is a more adequate account of the work done by instructional designers, teachers, and learners than his theory of technical human interests. This means that instructional designers will have to reorient their efforts in the following manner:

1. since all people bring preunderstandings to a situation and construct meanings in a situation, instructional designers will have to find ways to support this process rather than believe they can pre-specify learning outcomes for the learner and pre-specify instructional plans for the teacher.
2. since all people exercise practical judgement in the process of constructing meanings while teaching and learning, instructional designers will have to find ways to create useful resources for teachers and learners that support the meaning-making process. For example, instructional designers can pose problems and critical questions for teachers and learners rather than present predigested puzzles.
3. since right action in a given situation cannot be prespecified but only worked out by the participants, instructional designers will have to create learning resources and learning environments that leave some space for teachers and learners to work out their own sense of the good. Hence, instructional designers will have to give up the notion of designing 'teacher-proof' instruction and 'idiot-proof' learning resources.
4. since reflection and deliberation by the participants in a learning situation are so crucial to the creation of meaningful practice, instructional designers will have to find ways to avoid conceptualizing everything as a skill. This even applies to high-level abilities such as problem-solving, collaboration, and communication. Hence, instructional designers should not see problem-solving skills, or

collaboration skills, or communication skills as the highest form of learned capabilities but see these things as by-products of problem-posing judgments, collective deliberation, and collective meaning making.

5. since meaningful practice (be it instructional design practice, instructional practice, or learning practice) requires the negotiation of meaning as well as the negotiation of the terms of evaluation and power relations, instructional designers will have to go beyond the technological metaphor (i.e., beyond instructional plans and objectives, beyond learning theories) and participate directly in the teaching and learning experience. Barring this, they should at least create learning resources and learning environments that permit teachers and learners to construct knowledge as they see fit.

The shift that is implied in all of these claims is from a traditional instructional designer whose knowledges and practices are shaped by technological human interests to a learning resource designer and a learning environment designer whose knowledges and practices are shaped by a practical human interest.

Figure 8 About Here

BIBLIOGRAPHY

Apple, M. W. (1975). The adequacy of systems management procedures in education. In R.H. Smith (Ed.), Regaining educational leadership. New York: John Wiley.

Apple, M. W. (1979). Ideology and curriculum. London: Routeledge & Kegan Paul.

Apple, M. W. (1982). Teachers and texts. New York: Routeledge & Kegan Paul.

Brown, J. S. (1988). Steps toward a new epistemology of situated learning. Proceedings of the ITS-88. International Conference on Intelligent Tutoring Systems. University of Montreal. Montreal, Canada. June 1-3.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(1), 32-42.

Bullough, R. V., Goldstein, S. L., & Holt, L. (1984). Human interests in the curriculum: Teaching and learning in a technological society. New York: Teachers College.

Dreyfus, H. L., & Dreyfus, S.E. (1986). Mind over machine: The power of human intuition and expertise in the era of the computer. New York: Free Press.

Freire, P. (1970). Pedagogy of the oppressed. New York: Seabury Press.

Gagne, R. M. (Ed.). (1987). Instructional technology: Foundations. Hillsdale, NJ: Lawrence Erlbaum.

Gagne, R. M., Briggs, L.J., & Wager, W.W. (1988). Principles of instructional design. Third edition. New York: Holt, Reinhart, & Winston.

Grundy, S. (1987). Curriculum: Product or praxis? New York: Falmer Press.

Guba, E. G., & Lincoln, Y. S. (1989). Fourth generation evaluation. London: SAGE.

Habermas, J. (1972). Knowledge and human interests. Second edition. London: Heinemann.

Habermas, J. (1984). The theory of communicative action: Volume 1: Reason and the rationalization of society. Boston, MA: Beacon Press.

Habermas, J. (1987). The theory of communicative action: Volume 2: Lifeworld and system: A critique of functionalist reason. Boston, MA: Beacon Press.

Heinich, R. (1988). The use of computers in education: A response to Streibel. Educational Communications and Technology Journal, 36(3), 147-152.

Kliebard, H. (1987). The struggle for the american curriculum. New York: Routeledge, Chapman, & Hall.

Koetting, J. R. (1979). Towards a synthesis of a theory of knowledge and human interests, educational technology, and emancipatory education: A preliminary theoretical investigation and critique. Ph.D. Dissertation. University of Wisconsin, Madison.

Lave, J. (1988). Cognition in practice. Boston, MA: Cambridge University Press.

Livingston, D.W., & Contributers. (Eds). (1987). Critical pedagogy & cultural power. South Hadley, MA: Bergin & Garvey.

Nichols, R. (1989). Reconciling educational technology with the lifeworld: A study of Habermas' theory of communicative action. Proceedings of selected research presentations at the 1989 Annual Convention of the Association for Educational Communications and Technology, 341-355. (ERIC Document Reproduction Service No. ED 308805).

Nunan, T. (1983). Countering educational design. New York: Nichols Publishing.

Polanyi, M. (1958). Personal knowledge: Towards a post-critical philosophy. Chicago: University of Chicago Press.

Reigeluth, C. M. (Ed.). (1983). Instructional-design theories and models. Hillsdale, NJ: Lawrence Erlbaum.

Reigeluth, C. M. (Ed.). (1987). Instructional theories in action. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rogoff, B., & Lave, J. (Eds.). (1984). Everyday cognition: Its development in social context.

Shor, I. (1980). Critical teaching and everyday life. Boston, MA: South End Press.

Streibel, M. J. (1986). A critical analysis of the use of computers in the education. Educational Communications and Technology Journal, 34(3), 137-161.

Streibel, M. J., Stewart, J. H., Koedinger, K., Collins, A., & Jungck, J. (1987). MENDEL: An intelligent computer tutoring system for genetics problem solving, conjecturing, and understanding. Machine-Mediated Learning, 2(1&2), 129-159.

Streibel, M. J. (1988). A response to Robert Heinich and Suzanne Damarin. Educational Communications and Technology Journal, 36(3), 153-160.

Streibel, M. J. (1989). Instructional plans and situated learning: The challenge of Suchman's theory of situated action for instructional designers and instructional systems. Journal of Visual Literacy, 9(2), 8-34.

Suchman, L. A. (1987). Plans and situated actions: The problem of human/machine

Technical Approach to Instructional Design, Instruction, and Learning

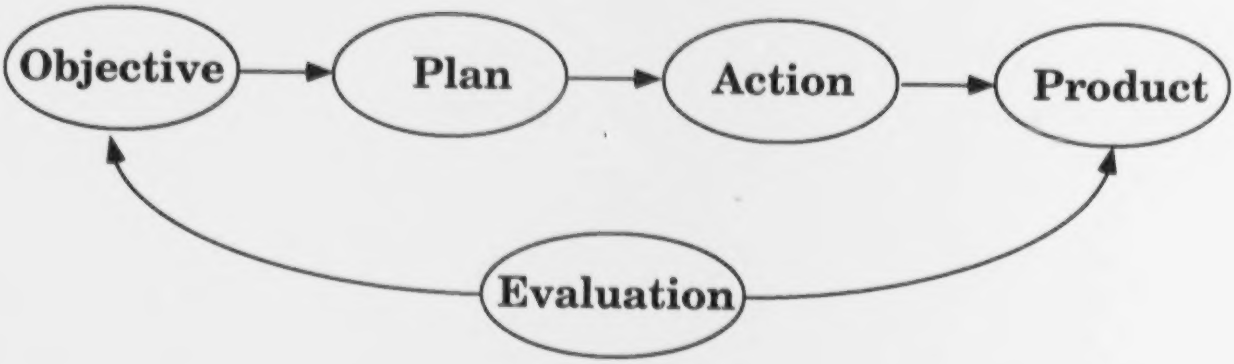

Objective
Plan
Action
Product
Evaluation

Figure 1

Practical Approach to Instructional Design, Instruction, and Learning

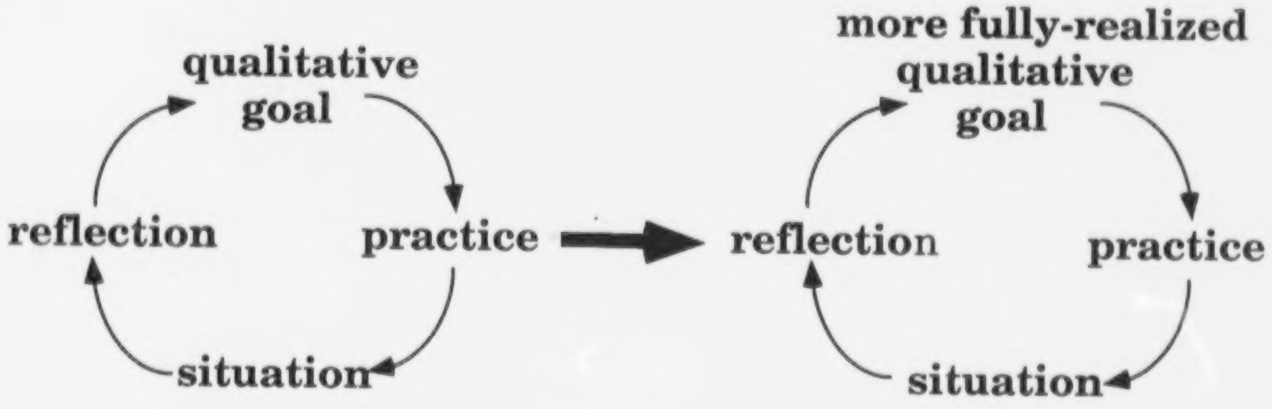

qualitative goal
reflection
practice
situation
more fully-realized qualitative goal
reflection
practice
situation

Figure 2

Social Factors in the Technical Approach

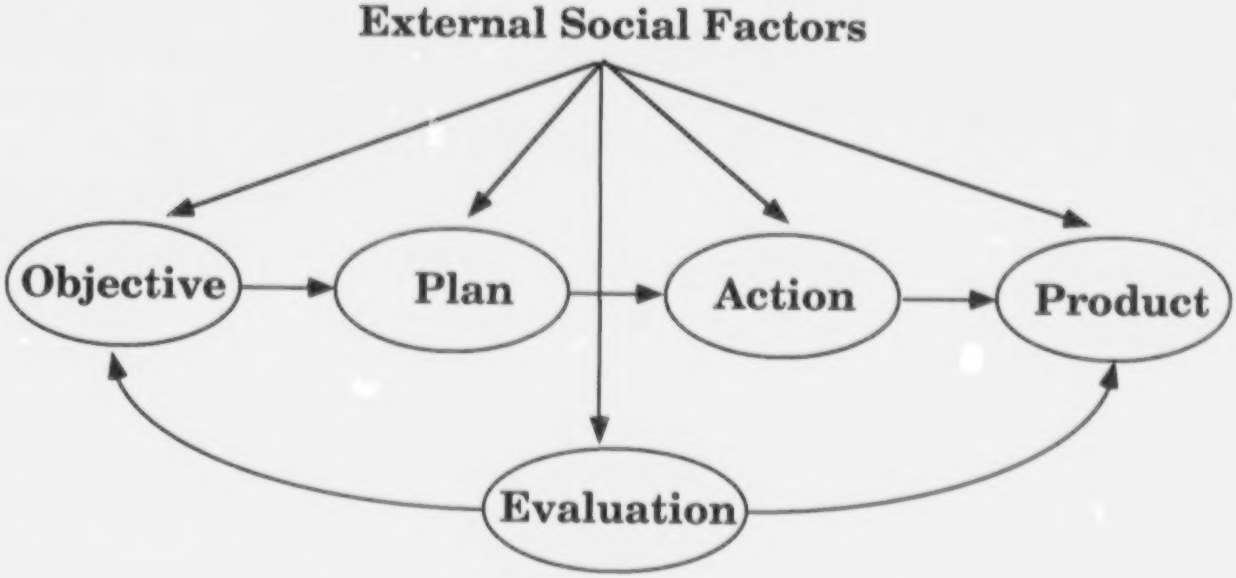

External Social Factors
Objective
Plan
Action
Product
Evaluation

Figure 3a

Social Component of the Practical Approach

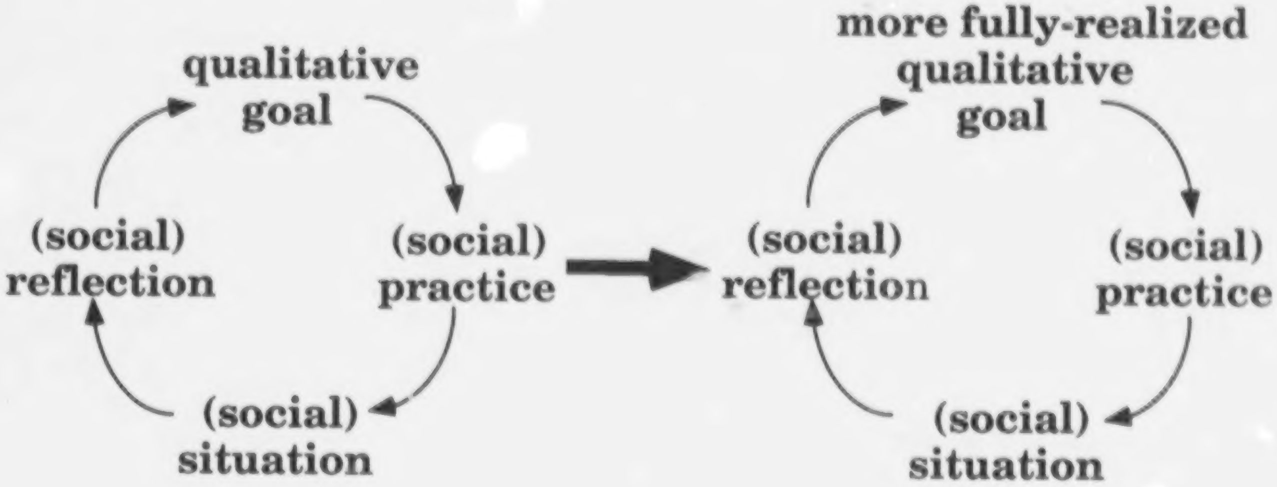

qualitative goal
(social) practice
(social) situation
(social) reflection
more fully-realized qualitative goal
(social) practice
(social) situation
(social) reflection

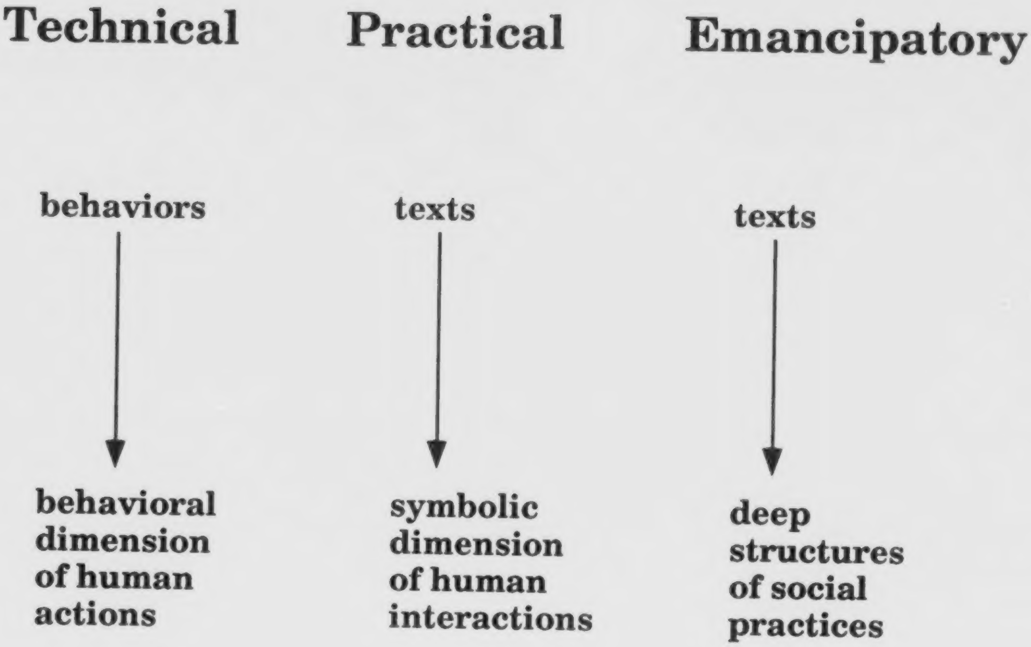
Figure 3b

Types of Human Interest

	Technical	Practical	Emancipatory
Ways of Knowing	empirical/ analytical	historical/ hermeneutic	critical/ cultural
Knowledge	facts, laws, procedures	narrative stories	critical theorems
Role of Theory	theory guides actions	theories as resources for actions	theorems help construct new knowledge and practice

Figure 4

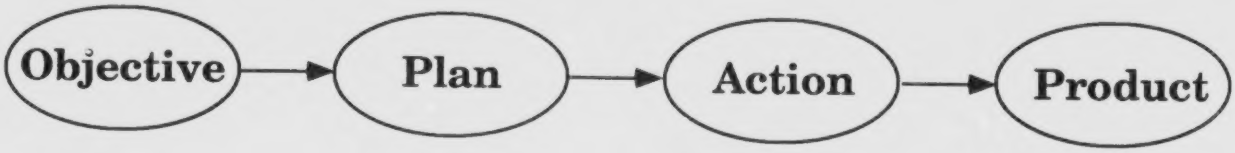
Objects of Study



Technical	Practical	Emancipatory
behaviors	texts	texts
↓	↓	↓
behavioral dimension of human actions	symbolic dimension of human interactions	deep structures of social practices

Figure 5

Skills Needed in the Technical Approach



Objective	Plan	Action	Product
- needs analysis	- design	- implementation	- measurement
- task analysis	- development	- management	- evaluation
- behavioral objectives	- planning		

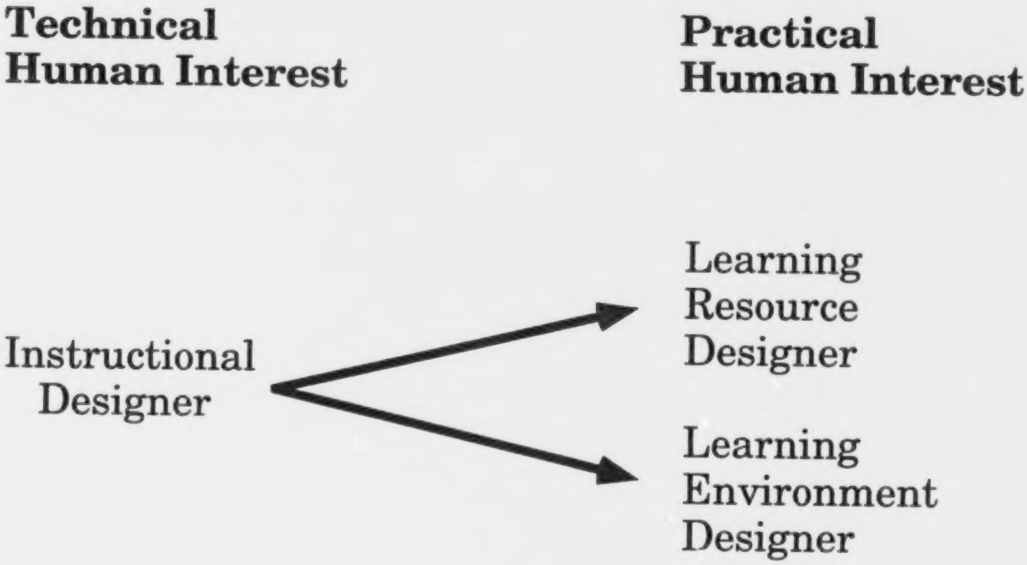
Figure 6

Dispositions in the Various Types of Human Interest

	Technical	Practical	Emancipatory
Basic Orientation	controlling self, other, and environment for external purposes	understanding self, other, and environment through interaction	restructuring social practices for justice and freedom
Outcomes	correct behavior	meaningful action	just society
Authority Resides In	plan	practitioner	historical community
Forms of Logic	instrumental logic	consensual logic	dialectical logic

Figure 7

Proposed Shift



Technical
Human Interest

Practical
Human Interest

Instructional
Designer

Learning
Resource
Designer

Learning
Environment
Designer

Figure 8

Title:

Automating Instructional Systems Development

Authors:

Robert D. Tennyson
Dean L. Christensen

Automating
Instructional Systems Development

Abstract

This paper presents framework specifications for an instructional systems development (ISD) expert system. The goal of the proposed ISD expert system is to improve the means by which educators design, produce, and evaluate the instructional development process. In the past several decades, research and theory development in the fields of instructional technology and cognitive science has advanced the knowledge base for instructional design theory such that learning and thinking can be significantly improved by direct instructional intervention. Unfortunately, these advancements have increased the complexity of employing instructional design theory, making instructional development both costly and time consuming. We are proposing that through the application of expert system methods, it is now possible to develop an intelligent computer-based ISD expert system that will enable educators to employ instructional design theory for curricular and instructional development. Presented in this paper is a framework for the development of an ISD expert system that will assist both experienced and inexperienced instructional developers in applying advanced instructional design theory.

Automating
Instructional Systems Development

Advancements in cognitive psychology and instructional technology in the past three decades have aided in the building of a literature of instructional design theory that can provide educators with sophisticated means to improve learning in all levels and conditions of education and training (Tennyson, 1990d). However, with this theoretical growth in instructional design has come the problem of instructional developers (i.e., any educator producing teacher-independent instruction) learning how to apply the new knowledge.

In response to this growth in the field of instructional design theory and practice, universities have developed graduate programs to produce instructional design (ID) experts. Even at the masters' level, these graduate programs require at least two years of full-time study. Therefore, if the educational community is to employ this body of knowledge to improve learning, it must either (a) develop inservice training programs to teach instructional design theory and practice or (b) develop a means by which educators can employ the knowledge without necessarily having to become ID experts.

This rapid growth in the instructional design field has also occurred in hundreds of other technical fields. To help maintain high levels of sophistication and to bring into application the most advanced knowledge from their respective fields, many of these other fields have employed expert system methods. An expert system is a computer-based representation of the domain-specific knowledge of an expert in a form that can be accessed by others for assistance in problem solving and decision making. An implication of this definition is that an inexperienced person can with the aid of an expert system perform tasks that would normally require the direct involvement of a domain-expert. Proposed in this paper are framework specifications for the development of an expert system to help instructional developers (i.e., authors) use the most advanced knowledge in the field of instructional design theory when designing and producing curriculum and instruction.

Instructional Systems Development

The process of designing, producing, and evaluating instruction is referred to in the literature as instructional systems development (ISD). The main components of ISD include (a) analysis of the instructional (and/or curricular) problem/need, (b) design of specifications to solve the problem, (c) production of the instruction, (d) implementation of the instruction, and (e) maintenance of the instruction. Embedded in each component of ISD are specific types of evaluation to insure quality control (Tennyson, 1978).

There are in the current literature several examples of computer-based tools intended to improve the productivity of the ISD process. Hermanns (1990) describes Computer-Aided Analysis (CAA), a computer program which aids in job task analysis. Based on a hierarchically-organized list of job tasks entered by the instructional designer, CAA produces as output a set of preliminary terminal learning objectives that can be further reviewed and edited by the developer. Ranker and Doucet (1990) describe SOCRATES, which allows the user to fill in information that is used by SOCRATES to create an instructor's lesson outline including objectives, events of instruction, samples of student behavior and test questions. Perez and Seidel (1990) present an overview of their specifications for an automated training development environment that will be based on the Army Systems Approach to Training (SAT) model of instructional design. The main features of the environment are a set of tools for developing the components of instruction and an expert design guide for assisting the designer in using the tools. Merrill and Li (1990) propose ID Expert, a prototype rule-based expert system for instructional development that makes recommendations about content structure, course organization, and instructional transactions (tutor/student interactions) based on information supplied by the designer.

The systems just referred to differ greatly in function and scope and in the degrees to which each makes use of expert system methods to reduce the level of knowledge required of instructional designers. This paper proposes framework specifications for an ISD expert system that would employ intelligent interface techniques to allow even the most inexperienced author to immediately begin to develop quality instruction. Labeled ISD Expert, the proposed system would make expert knowledge about the most sophisticated ISD methods readily available to potential authors, thus minimizing or eliminating the need for formal instructional systems development training.

The ISD model proposed for ISD Expert (see Figure 1) was developed to reflect an application model rather than a teaching model. That is, most ISD models are based on learning ISD, thus they resemble a linear process that attempts to include all possible variables and conditions of ISD. The result is that they do not take into account any other ISD situations other than complete start to finish instructional development. The assumption is that in all ISD situations, ISD starts at the analysis phase and proceeds step-by-step to the final completion of the implementation phase. The proposed ISD model, in contrast, views the author's situation as the beginning point of any possible ISD activities. For ISD Expert, the proposed ISD model is an associative network of variables and conditions, that can be addressed at any point in instructional development depending on the given situation.

This paper does not provide complete specifications for ISD Expert: instead, it provides a framework from which specifications can be designed and developed. The content of our ISD Expert includes both the philosophy of the proposed ISD Expert and the framework specifications. Given the complexity of ISD and the effort necessary to develop an expert system, hopefully, this chapter will also serve as a means for extending the dialogue on the concept of automated ISD systems and tools (e.g., see Merrill, 1990).

Philosophy of ISD Expert

Expert systems are designed for domain experts to aid them in dealing with complex processes that are either time consuming or which they do not have specific experience (e.g., in a sub-domain). In practice, expert systems have been successful when the content is narrowly focused and when the situations have clear rules for decision making (Smith, 1984). Because of the range of experience and training in ISD among instructional developers, I am proposing an expert system that will be designed for authors who are content domain experts but not necessarily ISD domain experts. This is not a contradiction of previous expert system efforts, but a reflection of the fact that the user of ISD Expert will not be an ISD expert initially; rather, the proposed ISD Expert would take into account a range of expertise and experience in instructional design theory and practice.

To accomplish this goal, we are further proposing an expert system that would employ intelligent human-computer interface techniques. The intelligent ISD Expert would operate at two basic levels: First, a coaching expert that would direct inexperienced authors through the acquisition of ISD skills while helping them deal with their specific situation; and, second, an advising expert that would assist experienced authors by making recommendations for their specific situation. For example, for an inexperienced author, the coaching function would deal with basic ISD skills and direct the development effort. In contrast, ISD Expert would function as an advisor to an experienced author, making recommendations while the author controlled the actual ISD decision making. In this environment, both inexperienced and experienced authors will be exposed to opportunities to increase their individual expertise through a process of learning ISD while using the system (Schiele & Green, 1990).

The importance of the distinction between the coaching and advisement functions is based on a review of research findings in expert systems. An example from this body of research is Clancey's (1979; 1983) work with MYCIN, a medical diagnosis consultant program, and GUIDON, a tutorial program designed to make use of MYCIN's rule base for teaching purposes. Clancey found that the rules encoded in MYCIN were inadequate for teaching because the knowledge required for justifying a rule and explaining an approach was lacking. He found it necessary to add additional components to GUIDON to help organize and explain the rules (Clancey, 1989). In a similar fashion, ISD Expert will have the ability to support and explain its recommendations and prescriptions in the language of ISD, not merely by enumerating the rules applied to make a recommendation. An example of one approach to providing this ability can be found in Swartout (1983). Swartout combined declarative and procedural knowledge, in the form of domain principles, to create the knowledge base for XPLAIN, a drug prescription consultant which provides detailed justification of its prescriptions.

Although ISD Expert can not be considered a means for teaching ISD, the very nature of the system's philosophy which assumes that authors will gain knowledge with experience, will result in continuing improvements in ISD applications. That is, as authors gain experience in ISD, the system would exhibit the characteristics of a conventional expert system. Therefore, it should increase the efficiency of instructional development and help in those areas where even experienced ISD authors initially lack specific expertise.

ISD Expert intelligent author-computer interface. ISD Expert, as proposed, would operate as an expert system employing intelligent author-computer interface (ACI) methods between the author and the system (Anderson, 1988). The ISD Expert intelligent ACI model (see Figure 3) would consist of four modules: the author's model of instruction, an ISD tutor (with both coach and advisor capabilities), an ISD knowledge base, and an instructional content knowledge base. Both knowledge bases would have knowledge acquisition capabilities. The ISD Expert tutor would be responsible for the interface between the individual authors and specific activities associated with developing their respective instructional needs.

ISD Expert system. In addition to the intelligent ACI component, the ISD Expert system would have three functions (see Figure 2). The first function would be to aid in the diagnosis of a given author's situation. This diagnostic function would evaluate the current situational condition(s) of the author (e.g., does the author want to prepare a computer-based graphic program for use in a lecture; does the author want to design a new course?). Following the situational evaluation, the second function would recommend prescription(s) along the lines associated with the level of author experience. That is, instead of trying to force all situations into a single solution, the prescription(s) would be individualized, based on situational differences and ISD experiences of the author. And with the third function, ISD Expert, through the system's tutor, would help the authors in accomplishing the prescriptions. As authors become increasingly more sophisticated in using ISD

Expert, they will be ready to accept increasingly more advanced variables and conditions of instructional design theory and practice.

ISD model. The proposed content for ISD Expert is the fourth generation ISD model (Tennyson, 1990). This ISD model is designed to adjust to future growth in instructional design theory and therefore does not become obsolete as new advancements are made, unlike earlier models. Figure 1 presents an illustration of the fourth generation ISD model. Briefly, the four generations of ISD models can be described as follows:

-First generation (ISD 1, Figure 1, Appendix A). The main focus of the first generation model was the implementation of the behavioral paradigm of learning (Glaser, 1966). The system consisted of four components: objectives, pretest, instruction, and posttest. The system was completed with an evaluation loop for purposes of revision.

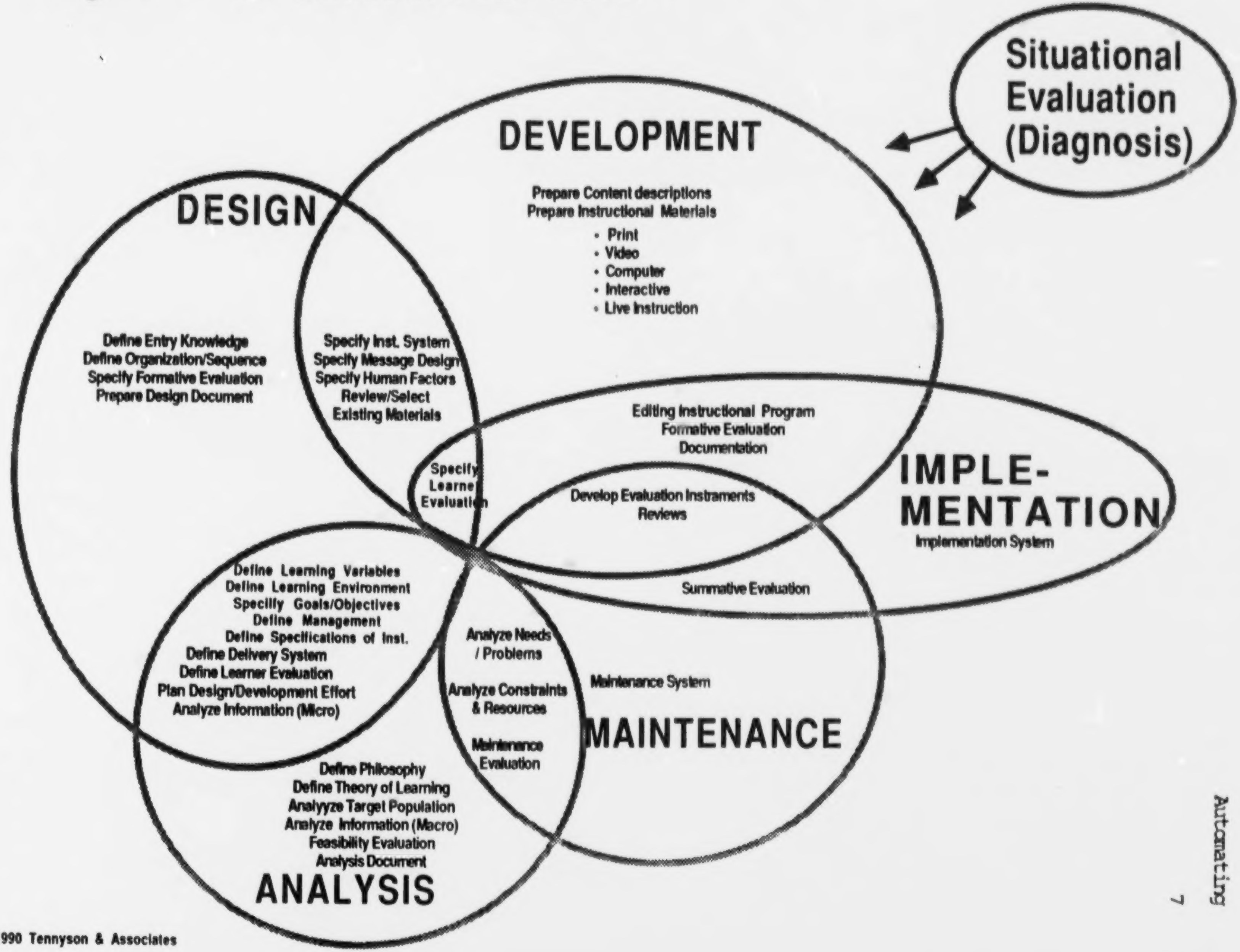
-Second generation (ISD 2, Figure 2, Appendix A). Advancements in instructional technology led to the need to increase the variables and conditions of the ISD model. The second generation adopted systems theory to control and manage the increasingly complex ISD process (Branson et al., 1976). The behavioral learning paradigm remained, but was of secondary importance to the focus of the system: developing instruction.

-Third generation (ISD 3, Figure 3, Appendix A). In practice, the ISD process was too linear and did not account for situational differences among applications (Tennyson, 1977). To account for situational differences, the external control of the system (i.e., the boxes and arrows) gave way to phases of ISD, that could be manipulated in any order by the instructional designer. This model assumed that ISD was an iterative process that could be entered at any point depending on the current state of the author's situation (Allen, 1986). Learning theory was still considered behavioral but cognitive theory was making some appearances (e.g., use of simulations for acquisition of cognitive skills in decision making).

-Fourth generation (ISD 4, see Figure 1). Advancements in cognitive psychology have made major changes in many of the ISD variables (e.g., content analysis, objectives, measurement, instructional strategies), making the ISD model yet more complex (see Tennyson, 1990a,b; Tennyson & Rasch, 1990). Employing technological developments from the field of artificial intelligence, the fourth generation model handles the complexity of ISD with a diagnostic/prescriptive system. Extending from the second and third generations, the ISD 4 model provides the knowledge base for the proposed ISD Expert system.

Cognitive theory. With growth of research and theory in cognitive psychology (Bonner, 1988), ISD Expert will exhibit a strong cognitive learning theory basis in both its ISD content and its approach to author-computer interaction. Early ISD models had a strong behavioral paradigm as their

Figure 1 4th Generation ISD Model



Situational Evaluation (Diagnosis)

DEVELOPMENT

Prepare Content descriptions
Prepare Instructional Materials

- Print
- Video
- Computer
- Interactive
- Live Instruction

DESIGN

Define Entry Knowledge
Define Organization/Sequence
Specify Formative Evaluation
Prepare Design Document

Specify Inst. System
Specify Message Design
Specify Human Factors
Review/Select
Existing Materials

Specify
Learner
Evaluation

Editing Instructional Program
Formative Evaluation
Documentation

Develop Evaluation Instraments
Reviews

IMPLE-MENTATION

Implementation System

Summative Evaluation

Define Learning Variables
Define Learning Environment
Specify Goals/Objectives
Define Management
Define Specifications of Inst.
Define Delivery System
Define Learner Evaluation
Plan Design/Development Effort
Analyze Information (Micro)

Analyze Needs / Problems

Analyze Constraints & Resources

Maintenance Evaluation

Maintenance System

MAINTENANCE

Define Philosophy
Define Theory of Learning
Analyyze Target Population
Analyze Information (Macro)
Feasibility Evaluation
Analysis Document

ANALYSIS

© 1990 Tennyson & Associates

learning theory foundation (e.g., the first generation ISD model). The instructional strategies embedded in the first generation ISD models followed closely the behavioral paradigm of small incremental steps with emphasis on reinforcement of correct responses. For example, a task analysis in the ISD 1 paradigm favored a sequential approach that included student exposure to all possible attributes in a domain of information. For the most part, situational context and higher order cognitive knowledge and strategies were not considered because they did not fit the behavioral paradigm that dealt only with temporal content and observable performances (Brown, Collins, & Duguid, 1989).

In many other aspects of instructional development, the first generation ISD models also incorporated the behavioral paradigm, especially in the evaluation of learners. The focus of learner evaluation was on attainment of performance objectives that were isolated from meaningful applications or situations. By the 1970s, however, the ISD models exhibited more characteristics of systems models, the result being a separation of ISD procedures from a given learning theory paradigm. This growth in the "systems" of ISD was referred to above as the second generation ISD model.

Although the learning theory foundation of the ISD 2 models remained basically behavioral, the inflexibility of the flow-chart nature of the models limited their utility. In response to this inflexibility, the ISD 3 models actually proposed the elimination of the linearity of ISD by including phases of development that could be manipulated according to the unique conditions of given situations. The third generation ISD model identified phases of ISD that included direct links to specific forms of evaluation (see Figure 3, Appendix A), and therefore allowed user control of the procedures based on situational need. The third generation focused on an increased emphasis on evaluation without basic changes in the learning theory foundation, although the flexibility of the model made it possible to include the growing literature in cognitive psychology.

By the end of the 1980s, there had been sufficient empirical and theoretical work in cognitive psychology and instructional technology to once again see the possible effects of learning theory on ISD (Glaser & Bassok, 1990). The effects can be seen in such things as the importance of macro (i.e., curricular) level activities in ISD, contextual analysis of the information to be learned, evaluation of the learners, employment of interactive media, instructional strategies for higher order thinking, employment of structured and discovery instructional methods, effect of the affective domain on the cognitive, influences of group interactions on learning, and context and situational variables on knowledge acquisition (Tennyson, 1990b). The result has been the development of fourth generation ISD models that resemble a schematic structure (see Figure 1) and have a cognitive learning paradigm foundation for the various procedures of instructional development. As stated earlier, Tennyson and Christensen's (in press) fourth generation ISD model is proposed for the knowledge base for ISD Expert.

Along with a cognitive learning foundation for the ISD content, we are proposing that the human-computer interface of ISD Expert exhibit a cognitive approach as opposed to a behavioral one. The contrast between the two approaches is the assumption made in regard to the interaction between the author and ISD Expert. In a behavioral approach the interaction between the author and the ISD expert system would be made at a reductionist level, that is, small incremental steps in linear sequence of instructional development in which the author is simply, and constantly, filling in requests for information without understanding the individual ISD tasks in relationship to the given situation. This is a common approach employed in expert systems for novices where the task is relatively concrete and the user is simply filling in information. However, it must be assumed that the ISD task is complex and requires an author who can intelligently use the system more productively as he/she gains experience. Therefore, a cognitive approach assumes, even initially, that the author can connect the individual ISD tasks with his/her given situation.

To summarize, a cognitive psychology implication for ISD Expert is an expert system that assumes that the author can from the start function in the role as an instructional designer. This implies that even at an initial level of ISD, the author will have a real instructional problem/need and that he/she will be able to solve the situation with the prescription(s) offered by the ISD Expert system. And, as the author becomes more experienced with ISD Expert, he/she will be able to make increasingly sophisticated use of the system. ISD is a complex process, but the complexity is in part due to the given situation. Thus, for the initial, inexperienced author, the potential employment of ISD Expert will focus on noncomplex situations, but with the author feeling that he/she is participating in real ISD decision making.

This approach to the author should limit training for ISD Expert to a set of basic software functions and activities. Instead of viewing training on ISD Expert in the conventional linear fashion where the author works through a set of meaningless practice situations, the training will be embedded in the initial individualized ISD situation. For example, if the author wants to develop a test, his/her initial entry into ISD Expert will deal with test construction. In other words, training and gaining experience will be driven by the individual author's situation. Rather than a two year graduate program as prerequisite to being an instructional designer, the author will be an instructional designer with ISD Expert beginning with his/her first time situation. Because ISD is a complex environment and the needs of individual authors will vary at any given time, over an extended period of time, the individual authors will acquire more ISD knowledge as situational needs occur.

Computer technology. Because ISD Expert is intended to improve the performance of instructional designers, rather than to advance the state of the art in expert systems techniques and methods, it is most productive to make use of existing, standard computer hardware and software architecture whenever possible in the development of ISD Expert.

Certain restraints are imposed on the hardware and software choices by the requirements of the environment in which ISD Expert will most often be applied. These requirements are summarized as follows:

- Support for several simultaneous authors at both local and remote sites;
- Large data storage capacity for knowledge bases and programs;
- Sophisticated graphics capability
- Provision for incorporating special-purpose programs (for example, to support research projects) into ISD Expert on an ad hoc basis;
- Employment of interactive media.

Where hardware is concerned, a basic decision is whether to implement ISD Expert on a central mainframe or minicomputer, or on microcomputers. Simons (1985) and Harmon, Maus, and Morrissey (1988) address the expanding role of the microcomputer in AI development, citing growing hardware capacity, wider availability of sophisticated software tools and increasing user familiarity with microcomputers as the forces contributing to the growth in expert system development for microcomputers.

We are proposing that ISD Expert be implemented in a network of PC's connected to a central network and file server with one or more large-capacity (perhaps 300 megabytes) hard disk drives for program and knowledge base storage. While there are a number of physical network topologies that could be used to implement ISD Expert, Figure 4 represents the general concept. There are some tradeoffs involved in using this configuration as contrasted with a network of "dumb" terminals connected to a single, central mainframe and data storage. For example, transmitting large quantities of data to/from the central file server to the PC's does require system overhead. However, the advantages outweigh the drawbacks. Given the local processing power of PC's, the intelligence of the system will be distributed throughout the system, minimizing the demands on the central unit. There is a large and growing quantity of AI software available for microcomputers at relatively low prices in contrast to mainframes. PC graphics are superior to all but the most sophisticated and expensive mainframe graphics systems.

The software used to create ISD Expert must provide an open architecture. That is, it must be practical to write local programs for special purpose functions (e.g., as research projects) and link them into the standard software with a minimum of effort. Also, the knowledge bases must be accessible to local programs as well as to the standard software. Expert system development is done either by using expert systems shells, which are commercially-available skeleton systems that can be instantiated with the specific domain knowledge required for an application, or by writing the expert system from scratch in a general or special purpose programming language. Harmon et. al. (1988) report that of 115 expert systems surveyed by them in actual use in the United States in 1986, 92 were produced using shells while 23 were written using programming languages (chiefly LISP).

Proposed is that ISD Expert be implemented using commercially-available expert system shells. However, in view of the fact that ISD Expert must also support customization, the shells that are chosen must support what are termed "own-code exits" to facilitate the linking in of custom programs. These custom programs must be written in a high-level computer language, preferably one with extensive AI features (e.g., LISP; PROLOG).

Summary

To establish a framework for ISD Expert, it is important to clearly specify the philosophy of the system (Morgan, 1989). A well specified philosophy will help keep the system under control during development and later when doing revisions. Proposed in this section is that ISD Expert have a foundation in cognitive psychology (Newell & Card, 1985). And, that this foundation specifies for the system both the content and the author-computer interface (Norman, 1986). Specific areas of the proposed philosophy are as follows:

-An expert system that has both diagnostic and prescriptive functions

-An expert system that will serve experienced and inexperienced authors

-An intelligent ACI system with both advising and coaching capabilities

-Knowledge base content will employ the fourth generation ISD model

-Employment of interactive media

-Cognitive learning theory as the foundation of the ISD procedures

-Cognitive paradigm approach to ACI

-Entry to system based on individual author situation

-Training as a concurrent activity with ISD activities

-A computer-based network system with remote capabilities

-Software tools that provide an open architecture

-Employment of a high-level language (e.g., an AI language)

-Commercial shells that include access to own-code programs

-Data dictionaries for knowledge acquisition components

The above discussion on a proposed philosophy for ISD Expert provides the foundation for the following section on framework specifications. The next section presents a basic framework for ISD Expert.

ISD Expert: System Framework

The purpose of this section is to propose framework specifications employing the above described philosophy of an expert system for instructional systems development. Because of the range in authors knowledge of ISD, we are proposing that ISD Expert be designed according to the methodology of intelligent human-computer interface systems. That is, rather than either attempting to teach ISD to the author or to develop a system around one linear approach that restricts and narrows the richness of ISD, our proposal is the design of a system that begins with the individual author's given situation. In this proposal, the intelligent ACI method will be concerned with improving both the authors application of ISD and their own models of instruction. As such, it will employ coaching and advising methods of human-computer interaction.

Furthermore, the proposed system will encourage the growth of the authors knowledge of ISD, but with the complexity of ISD being transparent. The purpose of ISD Expert will be to diagnosis the given situation of the author and then prescribe recommendations for dealing with his/her individual situation. It is assumed that each author will present a different situation and, therefore, will require a unique prescription. To accomplish this goal, the employment of heuristics is proposed for programming ISD Expert (see Bonnet, Haton, & Truong-Ngoc, 1988; Waterman, 1986). Two important features of the heuristic method, as contrasted, for example, with production rules, are (a) the flexibility needed to implement prescriptions in conditions of uncertainty or novelty (i.e., prescriptions are established in real time by integrating best available information from the system's knowledge base) and (b) the elimination of the need for an exhaustive reduction of ISD content knowledge to production rules.

One of the serious problems in expert systems design for nonstatistical areas has been the attempt to reduce complex and abstract concepts to production rules (e.g., Merrill's ID Expert, 1986). Even though we are proposing the use of a network and file server (with large capacity disk storage) system for the operation of ISD Expert, it is the programming time involved in trying to apply the reductionist approach to an environment as complex as ISD that rules out the exclusive use of the production rules programming methodology. The software architecture of ISD Expert must be open to allow for future extensions. The production rule method is not suitable for this type of complex situation (Clancey, 1983). So much of the ISD process is context bound; therefore, the system must be adaptable, allowing for prescriptions to be finalized by the author.

Proposed for ISD Expert is an expert system with four main components: an intelligent author-computer interface component, a diagnosis function component, a prescriptive function component, and an instructional production guide component (Figure 2). The intelligent ACI component will be the means by which authors will interact with ISD Expert. Rather than use a menu driven system, I am proposing a tutorial interaction between the author and ISD Expert. The diagnostic component will function as the evaluator of each author's situation and provide an evaluation report (Guba & Lincoln, 1986). This report will serve as the guidelines in preparing the prescription. Additionally, the prescription will be based on the author's ISD model as well as the diagnostic report. The fourth component will provide the author with assistance in the production of materials from the prescription(s). The level of assistance will again be influenced by the author's ISD model.

Intelligent Author-Computer Interface Component

The intelligent ACI component for the ISD Expert is illustrated in Figure 3. The main modules are as follows: (a) an author's model of ISD; (b) the ISD tutor; (c) the ISD knowledge base model; and (d) the content knowledge base. I will now discuss the role of each component of the ISD Expert tutor.

Author's model of ISD. The purpose of this module is twofold: (a) to establish the level of ISD expertise of the author and (b) to help the author improve his/her own model of instruction. This is necessary because no formal attempt is to be made to directly train the authors in ISD. The individual author's model will be updated with each use of the system. This profile of the author will help the system in its prescriptive recommendations. For example, experienced authors will have a narrow and limited knowledge of ISD and, also, of the ways in which their instruction could be improved; thus, prescriptions would be at their level of understanding. On the other hand, more experienced authors would be able to use more advanced prescriptions. It is important to keep the ISD prescriptions at the level of the author's experience and also to provide an opportunity for creativity and the possible use of different ideas generating from the author (Russell, Moran, & Jordan, 1988). A key feature of the proposed ISD Expert is the power of the author to disagree with a given prescription and still to be able to continue with the ISD process.

ISD tutor. Intelligent HCI systems work on the premise that a meaningful dialogue must be established between the user and the system. An important feature of the dialogue is the mixed initiative, where the user has an opportunity to query the system as well as being controlled by the direction of the system. The ISD Expert tutor will approach the diagnostic function from the context (situation) of the author. Personalizing the diagnostic activity will provide the opportunity for the tutor to search the content knowledge base to include specific references in the prescription to available existing materials and resources.

Because of the range of knowledge and experience in ISD of potential authors, two basic modes of interface are proposed. At one extreme will be authors who are completely inexperienced in ISD. For these individuals, a coaching mode is proposed. The coaching mode is a well established method of instruction used in intelligent computer-assisted instruction (ICAI). This

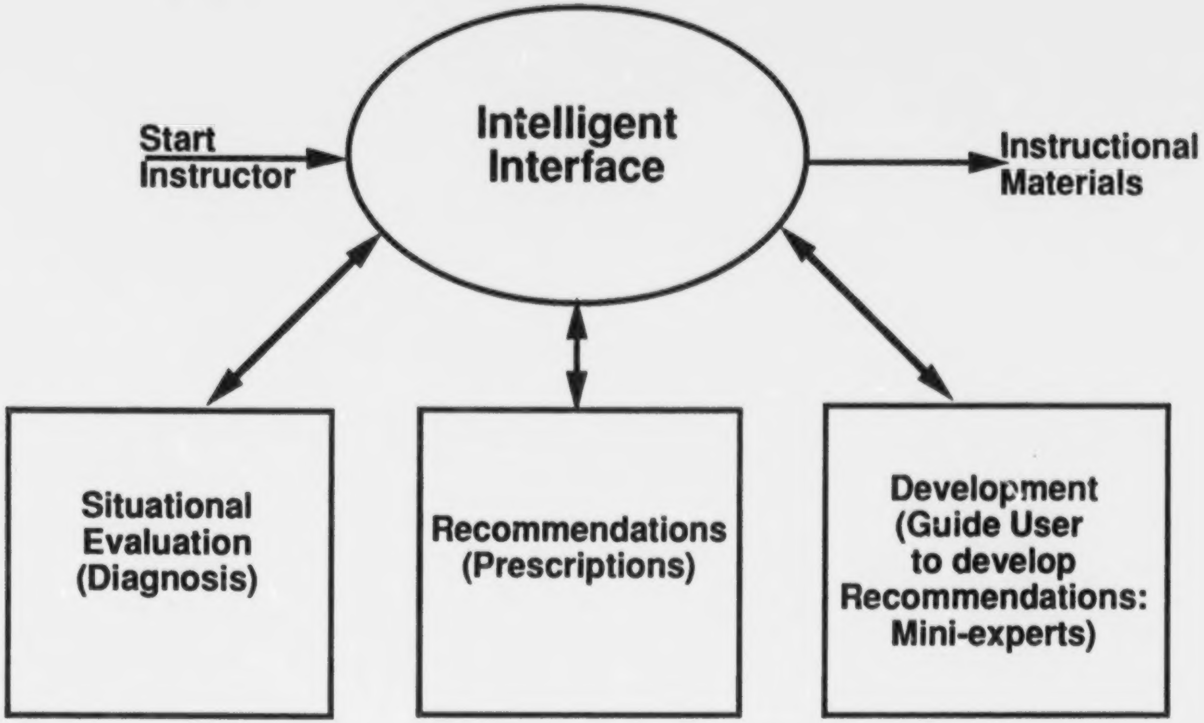

Start
Instructor
Intelligent
Interface
Instructional
Materials
Situational
Evaluation
(Diagnosis)
Recommendations
(Prescriptions)
Development
(Guide User
to develop
Recommendations:
Mini-experts)

Figure 2. ISD Expert Components.

© 1990 Tennyson & Associates

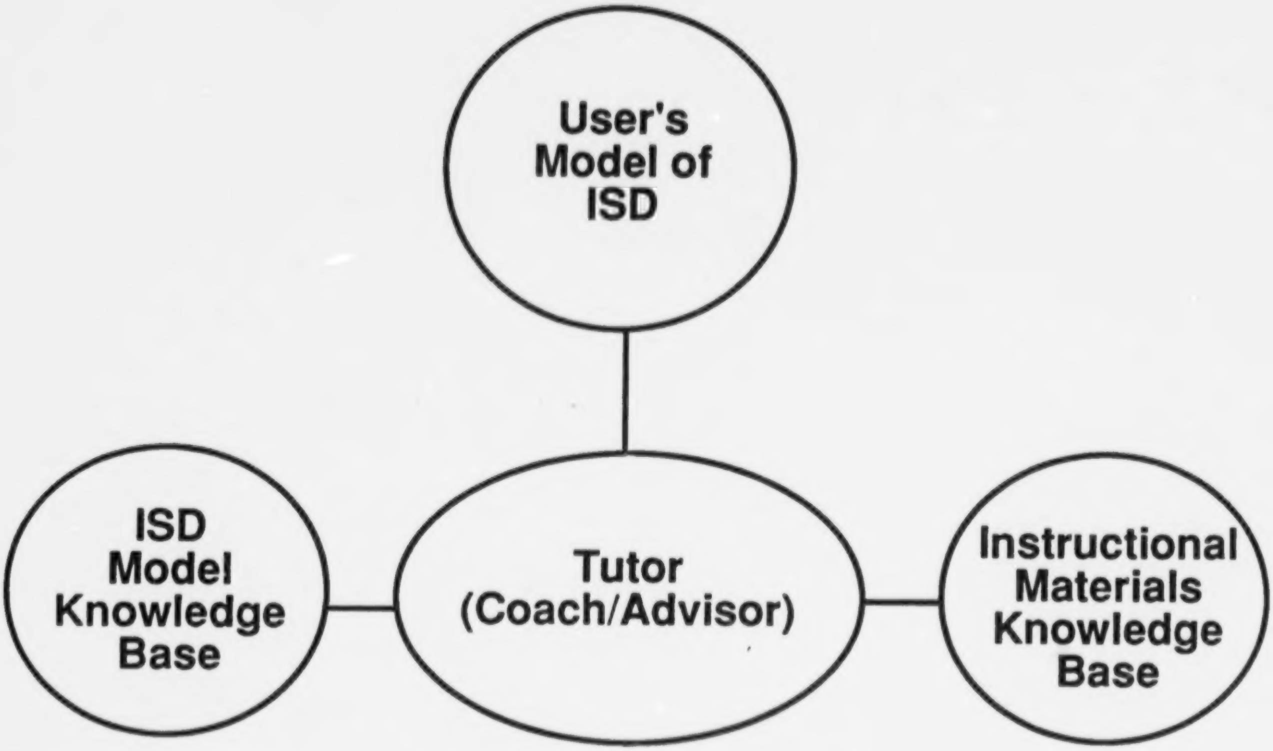

User's Model of ISD
ISD Model Knowledge Base
Tutor (Coach/Advisor)
Instructional Materials Knowledge Base

Figure 3. AIDA Intelligient Interface Component.

© 1990 Tennyson & Associates

mode assumes that the author will need direct and controlled assistance in dealing with his/her given situation. The function of the tutor as coach is to approach the ISD activity in a disciplined way while helping the author develop ISD skills. Prescriptions for the situation are specific and the coach is responsible for the decision making. In contrast is the advising interface. For the experienced author, the tutor as advisor would offer alternative prescriptions, with the final decision(s) in the hands of the author.

The tutor, as part of the intelligent ACI component, is the point of contact between the author and the other ISD Expert components (see Figure 2). In the proposed design, the tutor gathers information about the author's specific situation and, by interaction with the Situational Evaluation component, prepares a report of the given problem/need. This evaluation report is sent to the Recommendations component where a prescription(s) is prepared. When the prescription(s) is prepared, the tutor presents it to the author; at that point, depending on the mode of the tutorial interaction (i.e., coaching or advising), there may occur a dialogue between the author and the tutor to finalize the prescription. Once a final prescription is prepared, the tutor interacts with the ISD model knowledge base to set up the authoring activities. The tutor also assists the author in certain aspects of materials production through the fourth component of the ISD Expert system. Updating of the author's model will be the continuing role of the tutor in ISD Expert.

ISD model knowledge base. The content knowledge of ISD Expert will reside in the ISD model knowledge base (KB) (see Figure 1). Once the prescription(s) is decided upon, the necessary authoring activities are compiled by the tutor from the ISD model knowledge base and presented to the author. (Authoring activities of the knowledge base are presented in Appendix B.) Information within this KB will be stored as structured data files, organized as an associative network. The purpose here is to efficiently locate information without the restrictions of rigid production rules. That is, the ISD model knowledge base should exhibit the heuristic search characteristics of an information retrieval system.

Content knowlege base. The fourth module of the proposed intelligent ACI system for ISD Expert, the content knowledge base, is a source from which curricular and instructional materials resources may be obtained. These materials may be included in the implementation of prescriptions developed by ISD Expert or they may stand alone. For example, if an author wants a simulation for a given lecture, he/she could query the content knowledge base to see what might be available. In another situation, ISD Expert may develop a prescription and obtain the necessary materials from the content knowledge base without the author explicitly requesting the action. Access to the content knowledge base may be either by direct author query via the tutor or indirectly as a result of the implementation of prescriptions.

The content knowledge base will help eliminate duplication of effort in instructional development by providing a catalog of available materials.

Information in the content knowledge base would come from two sources. Material that is developed on ISD Expert as a result of instructional development can be added to the content knowledge base. Material may also be input from sources external to ISD Expert. For example, many materials and resources that are developed in R & D efforts independently of ISD would be useful in course applications if authors had access to them. General information manuals and other media-based resources (e.g., video disk materials) are another example of materials from external.

Situational Evaluation Component

The first activity in the proposed ISD Expert system is the evaluation of the given author's situation. The assumption is that each author will have a different need or problem, depending on his/her given situation. As the ISD Expert tutor establishes the author's model of instruction (see Figure 3), the Situational Evaluation Component will diagnosis the situation employing AI techniques. Again, it is assumed that the tutor will determine the experience level of the author and in turn adjust the report of the evaluation. For example, if the tutor determines that the author is experienced in ISD, and the situation is to develop a lesson on trouble shooting, the report would indicate those two conditions, which would influence the type of prescription(s) recommended. By focusing on the given situation, ISD Expert can employ the complexity and richness of the fourth generation ISD model without directly training the author about the entire model.

Recommendations Component

The purpose of ISD Expert is to help authors improve their instructional product development by applying the most advanced variables and conditions of instructional design theory. This is made possible by the recommendations component, which interacts with the ISD knowledge base to interpret the situational evaluation diagnosis and recommend a prescription to deal with the given instructional situation. Also, the prescription is adjusted to the author's level of experience. This is an important feature of the proposed ISD Expert because it prescribes an effort of development that can be efficiently accomplished by the author. For example, if an inexperienced author is presented with a prescription that would fit an experienced author's profile, the novice author would not be able to adequately follow the production activities (Component 4). The result would be that the prescription is implemented inefficiently or not at all. Presentation of the prescription will likewise be based on the experience of the author. The experience level of the author will determine the program control (i.e., coaching or advisement) employed in the production component.

Production Component

The term production is used here to reflect a variety of different types of instructional situations that might occur. ISD includes, in addition to instructional development, test development, computer-based management development, print materials development, instructional aids, visual aids, etc. The function of this component is to guide the author in the production process. As such, this part of the expert system directly interacts with the tutor. Because of the range of ISD activities, this component would be composed of mini-experts, each reflecting a different authoring activity (see Appendix B). That is, the mini-experts would be the various activities within the ISD model. For example, if the situation is to develop a test for trouble shooting, the author's model indicates an experienced author, and the prescription recommends a simulation, a mini-expert on design of simulations within component 4 would guide the author in the production of an appropriate simulation. An important feature of ISD Expert will be to facilitate the employment of avdvanced interative technology for instructional delivery. For example, for computer-based instruction, this component would directly produce the courseware (Tennyson, 1990c).

Once the production effort is completed, component 4 would send a report back to the tutor to update the author's model and to reference the effort in the content knowledge base. To further improve the efficiency of ISD Expert, instructional strategy (IS) shells will be accessible by the mini-experts to do the actual product development: IS shells would only require that the author enter into the system content information and the system would develop the product.

The above four components of the proposed ISD Expert system would be designed and programmed as independent expert systems so as to allow for future additions and elaborations. This is necessary because of the continuing growth in the instructional design theory field. That is, most expert systems are designed for specific, contemporary applications; when changes occur, a new expert system is designed and implemented.

Central Network System

In this section, we propose for the computer-based environment a configuration for a centrally-based network and file server for both local and remote PC workstations (Figure 4). Because of the proposal for a content knowledge base (see Figure 3) with acquisition capabilities and an author's instructional model within ISD Expert, a large capacity disk storage should be an integral part of the system. Also, given the computing power of PCs, much of the intelligent interfacing would take place at the workstation.

Although there are a large number of commercially-available shells for program development (e.g., HYPERcard), most do not allow for "own-code" exits. With such software, the development effort becomes constrained by the closed architecture of the given shell; the shell becomes a methodology in itself rather than a tool to be used in implementing multiple methodologies. I am proposing that ISD Expert be programmed in a high-level language with artificial intelligence features, but that commercially-available shells be used when feasible to augment the system features.

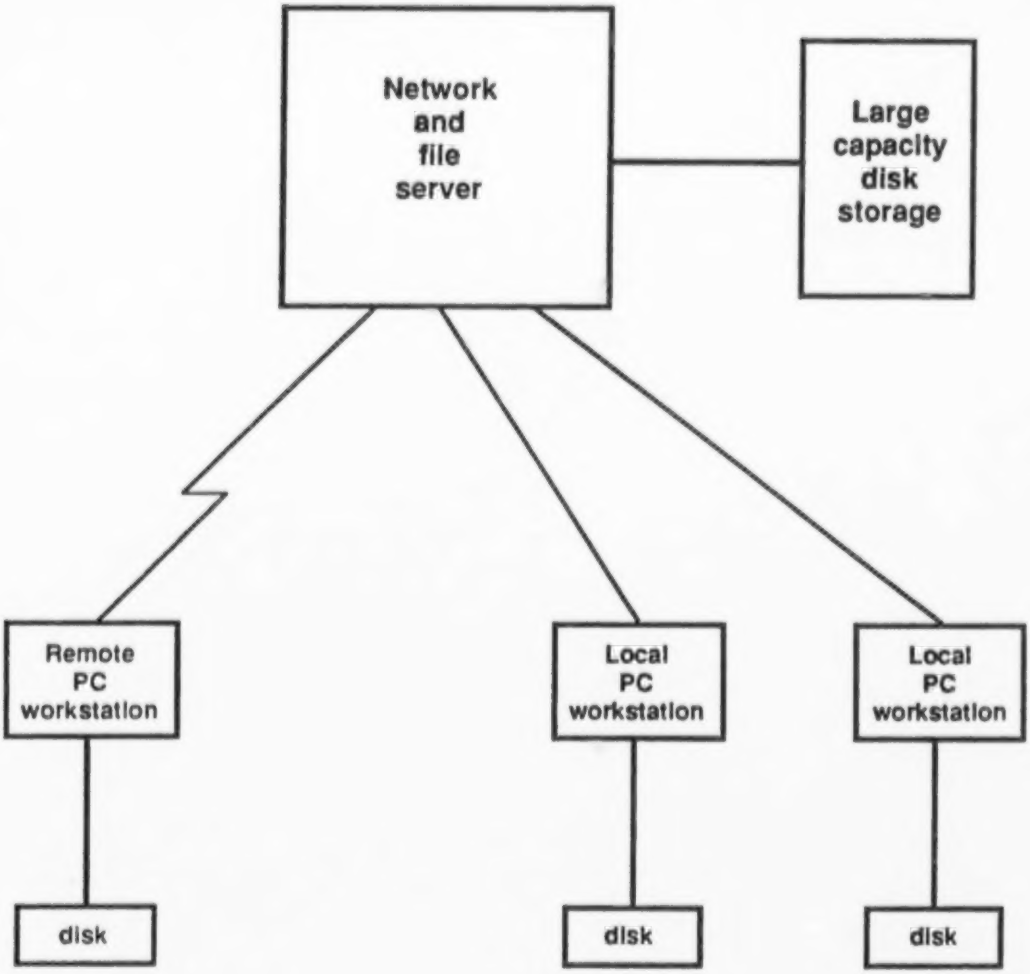

Network
and
file
server
Large
capacity
disk
storage
Remote
PC
workstation
Local
PC
workstation
Local
PC
workstation
disk
disk
disk

Figure 4. Distributed processing with PC workstations in local and remote locations connected to central file and network server.

Development Plan

The framework specifications presented in this paper offer a complete expert system for automating instructional development. To produce such a system, there are two possible approaches. The first would be to develop the ISD Expert system as presented. The second would be to follow an incremental approach in which an initial prototype is developed that only has a minimal set of features and is aimed at an experienced author. That is, an ISD Expert that would only have an advisor level tutor and the situational evaluation and recommendations components. The content knowledge base and acquisition features of the intelligent interface tutor and the production component would be added in subsequent elaborations.

Although the first approach seems possible, there are a number of problems that need to be considered that might favor the second approach. An initial problem is the cost factor. As stated earlier the majority of expert systems are developed using commercial shells. Cost in terms of software is the time required to produce a product that will be timely and profitable. That is, the proposed ISD Expert would most likely be a software product that would need to generate income within a reasonable timeframe. Rapid prototyping is a procedure to develop software employing shells that are linked by some general language (Hewett, 1989). Thus, instead of five years to produce a complete version of ISD Expert, an initial prototype could be developed in much less time.

A second major problem in producing a complete ISD Expert in the first approach is the necessary research needed for the new system. There has been minimal empirical research to date on instructional variables and conditions associated with the extension of cognitive learning theory to instructional design theory. Even though it is possible to develop an initial prototype, research in cognitive instructional design theory needs to be done as well as the interaction of media within this theoretical framework. A third problem area relates to the specification of the human-computer interaction variables and conditions necessary to run and manage the complex environment of ISD Expert.

Within the constraints defined above for approach one, we recommend the following incremental approach to ISD Expert development.

1. Framework specifications. This step conceptualizes the idea or vision of the expert system. This chapter serves as an example of the first activity in producing an automated instructional development system.

2. Functional specifications. From the initial outline of the basic system, the specific functions provided by the system need to be defined. From this step a rapid prototype can be developed as follows:

-Write functional specifications;

-Summarize what is known/not known about the functions;

-Estimate the complexity of the functions;

-Based on the summary and estimates, group the functions into ISD Expert 1 (i.e., a prototype), and then prioritize the functions for successive implementations for versions ISD Expert 2, ISD Expert 3, etc. Each version would add layers of functions and increased use of a high-level computer language.

3. Logical design. Starting with the prototype, define the logical components that provide the specified functions.

4. Physical design. Define the software modules which implement the logical design of the system.

5. Programming. With the prototype, rapid software development is recommended while with the successive versions, the software procedures defined in this paper would be followed.

6. Testing. Once the prototype is developed, it should be tested following standard computer software benchmark criteria.

7. Implementation. Complete the remaining tasks to implement ISD Expert1 while simultaneously accumulating the experience and research findings needed to produce ISD Expert 2.

8. Incremental development. Basically starting with number 3 above, iteratively build ISD Expert towards a system that includes all of the functions defined in numbers one and two.

Cycling through an incremental approach to development of ISD Expert would produce an initial product for employment and research within the constraints of costs and system knowledge. The initial prototype (or ISD Expert1) would exhibit many standard characteristics of the third generation ISD model (see Figure 3 Appendix A) with successive versions taking on more of the ideas associated with the fourth generation ISD model (see Figure 1).

Conclusion

The purpose of ISD Expert is to improve learning by aiding authors in the employment of contemporary instructional design theory. This paper presents framework specifications for an expert system to implement the concept of ISD Expert. Because of the experience range of authors for ISD Expert in terms of instructional design theory and practice, we are proposing an expert system that employs an intelligent human-computer interface system. That is, ISD Expert will interact with authors on an individual basis according to their respective experience with principles and variables of instructional design theory. ISD Expert will dialogue with authors along a continum of decision control ranging from system control (coaching function) to complete author control (advising function). Inexperienced authors will be coached to develop basic ISD skills while the more experienced authors will be advised on the employment of advanced instructional design variables and conditions.

An author's instructional design model is a necessary module for an intelligent human-computer interface component because it replaces the need for a separate training program for authors. The sophistication of ISD Expert's prescriptions will be directly influenced by the author's instructional design model. Therefore, both ISD novices and experts will be able immediately to use ISD Expert. That is, the proposed system would take into account experience in instructional design theory and practice.

Proposed for the ISD knowledge base is the fourth generation ISD model (see Figure 1). The initial set of authoring activities for the ISD knowledge base is presented in Appendix B. The content knowledge base is proposed as a data base for instructional materials within subject matter areas. Both knowledge bases will have acquisition capabilities.

The basic proposed ISD Expert system will have four interactive components: (a) an intelligent author-computer interface component; (b) a situational evaluation component (diagnosis), (c) a recommendations component (prescriptions), and (d) a production component. Proposed is that ISD Expert be designed for a computer-based network system using a high level AI type language and expert system shells. The system will also employ a large capacity disk storage for the two knowledge bases (i.e., ISD model and instructional content and materials); both knowledge bases will employ acquisition capabilities.

To implement the concept of the fourth generation ISD model, the situational evaluation component of ISD Expert will diagnose each author's given problem and/or need. This will make the system application orientated rather than the conventional lock-step system that is most suited for the teaching of ISD. From the diagnosis, ISD Expert will generate a prescription. For those authors who seek assistance in implementing the prescription, especially those requiring the development of instructional materials, the fourth component will guide the production effort. This production component will be composed of mini-expert systems that have specific functions (e.g., instructional strategy shells).

Because of both development costs and gaps in instructional design theory, we recommend an incremental approach to the development of ISD Expert. Initially, the project should begin with rapid prototyping techniques to produce a version one of ISD Expert. Subsequent versions would be elaborated according to the functional specifications as outlined in this chapter and from on-going research findings.

In conclusion, we are proposing an expert system that will bring the power of instructional design theory and practice to educators who would not normally have the opportunity to employ such knowledge in their instructional efforts. The proposed ISD Expert will improve learning by making instructional development both effective (i.e., by employing the most advanced principles and variables of learning and instructional theories) and efficient (i.e., by reducing the time and cost of conventional methods of instructional development).

References

Allen, M. W. (1986). Authoring aids: The means to improved authoring procedures. Proceedings of the SALT Conference. Orlando, FL: Society for Applied Learning Technology.

Anderson, J. R. (1988). The expert module. In M. C. Polson & J. J. Richardson (Eds.), Intelligent tutoring systems (pp. 21-53). Hillsdale, NJ: Erlbaum.

Bonner, J. (1988). Implications of cognitive theory for instructional design: Revisited. Educational Communication and Technology Journal, 36, 3-14.

Bonnet, A., Haton, J-P., and Truong-Ngoc, J-M. (1988). Expert systems: Principles and practice. Englewood Cliffs, NJ: Prentice Hall.

Branson, R., Rayner, G. T., Cox, J. L., Furman, J. P., King, FJ, & Hannum, W. H. (1976). Interservice procedures for instructional systems development (5 Vols.) (TRADOC Pam 350-30 and NAVEDTRA 106A). Fort Monroe, VA: U. S. Army Traning and Doctrine Command. (NTIS No. ADA-019 486 ADA-019 490)

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. Educational Researcher, 18(10, 32-42.

Clancey, W. J. (1979). Tutoring rules for guiding a case method dialogue. International Journal of Man-Machine Studies, 11, 25-49.

Clancey, W. J. (1983). The epistemology of a rule-based expert system—a framework for explanation. Artificial Intelligence, 20, 215-251.

Clancey, W. J. (1989). The knowledge level reinterpreted: Modeling how systems interact. Machine Learning, 4, 287-293.

Glaser, R. (1966). The design of instruction. In J. I. Goodlad (Ed.), The changing American school: NSSE 65th Yearbook (pp. 215-242).

Glaser, R., & Bassok, M. (1989). Learning theory and study of instruction. Annual Review of Psychology, 40, 631-666.

Guba, E. G., & Lincoln, Y. S. (1986). The countenances of fourth-generation evaluation: Description, judgment, and negotiation. Evaluation studies review annual, 11, 70-88.

Harmon, P., Maus, R., & Morrissey, W. (1988). Expert system tools and applications. New York: Wiley.

Hewett, T. T. (1989). Towards a rapid prototyping environment for interface design: Desirable features suggested by the electronic spreadsheet. In

A. Sutcliffe & L. Macaulay (Eds.), People and computers V (pp. 305-314). Cambridge, UK: Cambridge University Press.

Hermans, J. (1990). Computer-aided instructional system development. Educational Technology, 30(3), 42-45.

Merrill, M. D. (1986). Prescriptions for an authoring system. Journal of Computer-based Instruction, 12, 90-96.

Merrill, M. D. (1990). Introduction to special issue: Computer-based tools for instruction design. Educational Technology, 30(3), 5-7.

Merrill, M. D., & Li, Z. (1990). An instructional design expert system. In S. Dijkstra, B. van Hout Wolters, & P. C. van der Sijde (Eds.), Research on instruction: Design and effects (pp. 21-44). Englewood Cliffs, NJ: Educational Technology.

Morgan, C.C. (1989). Programming from specifications. New York: Prentice-Hall.

Newell, A., & Card, S. (1985). Prospects for psychological science in human computer interaction. Human Computer Interaction, 1, 209-242.

Norman, D.A. (1986). Cognitive engineering. In User-centered system design (pp.31-62). Hillsdale, NJ: Erlbaum.

Perez, R. S., & Seidel, R. J. (1990). Using artificial intelligence in education: Computer-based tools for instructional development. Educational Technology, 30(3), 51-58.

Ranker, R. A., & Doucet, R. M. (1990). SOCRATES: A computer-based lesson development advisor. Computers in Human Behavior, 6, 162-171.

Russell, D. Moran, T., & Jordan, D. (1988). The instructional design environment. In J. Psotka, L. Massey, & S. Mutter (Eds.), Intelligent tutoring systems: Lessons Learned (pp. 96-137). Hillsdale, NJ: Erlbaum.

Schiele, F., & Green, T. (1990). HCI formalisms and cognitive psychology: the case of Task-Action Grammar. In M. Harrison & H. Thimbleby (Eds.), Formal methods in human-computer interaction (pp. 9-62). Cambridge, UK: Cambridge University Press.

Simons, G. L. (1985). Expert systems and micros. Manchester, UK: NCC Publications.

Smith, S. F. (1984). Adaptive learning systems. In R. Forsyth (Ed.). Expert systems: Principles and case studies. New York: Chapman and Hall.

Swartout, W. R. (1983). XPLAIN: A system for creating and explaining expert consulting programs. Artifical Intelligence, 21, 285-325.

Tennyson, R. D. (1977). Instructional systems: Development, evaluation, and management of learning environments. In D. J. Treffinger, J. K. Davis, & R. E. Ripple (Eds.), Handbook on teaching educational psychology (pp. 215-242). New York: Academic Press.

Tennyson, R. D. (1978). Evaluation technology instructional development. Journal of Instructional Development, 2(1), 19-26.

Tennyson, R. D. (1989). Cognitive science and instructional technology updates of ISD. Technical Report. U.S. Air Force Human Resources Laboratory/IDC. Brooks Air Force Base, San Antonio, TX.

Tennyson, R. D. (1990a). A proposed cognitive paradigm of learning for educational technology. Educational Technology, 26(6), 12-17.

Tennyson, R. D. (1990b). Cognitive learning theory linked to instructional theory. Journal of Structural Learning, 13, 362- 371.

Tennyson, R. D. (1990c). Computer-based enhancements for the improvement of learning. In S. Dijkstra, B.H.A.M. van Hout Wolters, & P. C. van der Sijde (Eds.), Research on instruction: Design and effects. Englewood Cliffs, NJ: Educational Technology.

Tennyson, R. D. (1990d). Integrated instructional design theory: Advancements from cognitive science and instructional technology. Educational Technology, 26(6), 18-28.

Tennyson, R. D., & Rasch, M. (1988). Linking cognitive learning theory to instructional prescriptions. Instructional Science, 17, 369-385.

Tennyson, R. D., & Rasch, M. (1990). Instructional design for the improvement of learning and cognition. In H. Feger (Ed.), Wissenschaft und Verantwortung (pp. 71-90). Goettingen, Fed. Rep. of Germany: Hogrefe-Verlag.

Waterman, D. A. (1986). A guide to expert systems. Reading, MA: Addison-Wesley.

Appendix A

Figure 1. 1st Generation ISD Model (1960s).

Figure 2. 2nd Generation ISD Model (1970s).

Figure 3. 3rd Generation ISD Model (1980s).

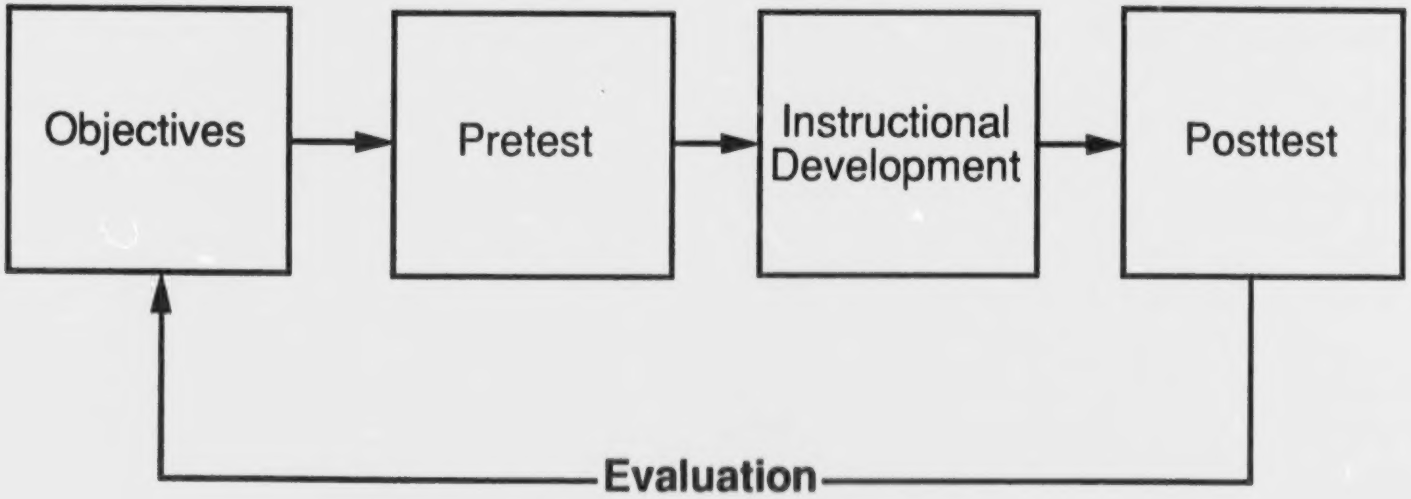

Objectives
Pretest
Instructional Development
Posttest
Evaluation

Figure 1. 1st Generation ISD Model (1960s).

© 1990 Tennyson & Associates

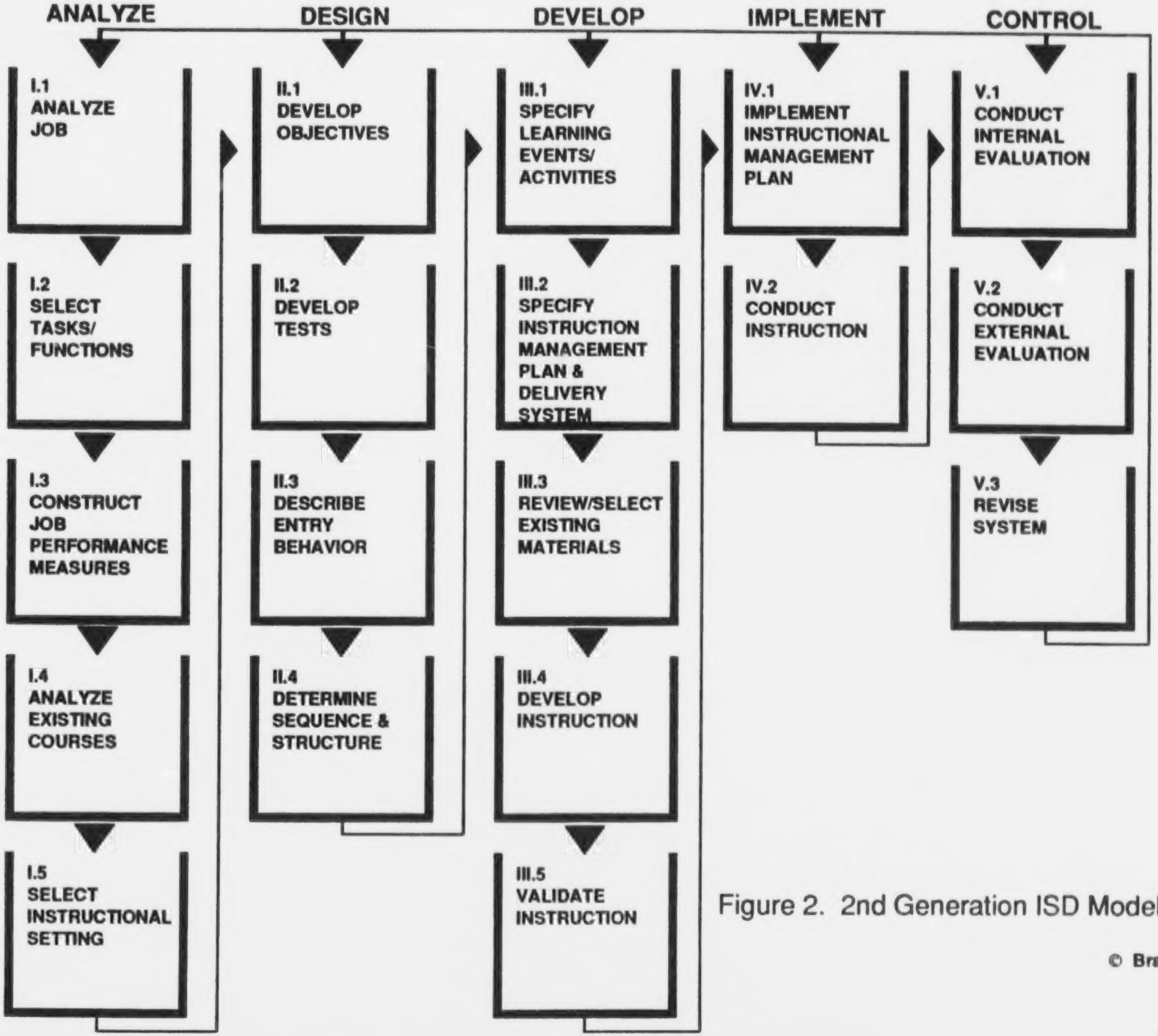

ANALYZE
DESIGN
DEVELOP
IMPLEMENT
CONTROL
I.1 ANALYZE JOB
I.2 SELECT TASKS/ FUNCTIONS
I.3 CONSTRUCT JOB PERFORMANCE MEASURES
I.4 ANALYZE EXISTING COURSES
I.5 SELECT INSTRUCTIONAL SETTING
II.1 DEVELOP OBJECTIVES
II.2 DEVELOP TESTS
II.3 DESCRIBE ENTRY BEHAVIOR
II.4 DETERMINE SEQUENCE & STRUCTURE
III.1 SPECIFY LEARNING EVENTS/ ACTIVITIES
III.2 SPECIFY INSTRUCTION MANAGEMENT PLAN & DELIVERY SYSTEM
III.3 REVIEW/SELECT EXISTING MATERIALS
III.4 DEVELOP INSTRUCTION
III.5 VALIDATE INSTRUCTION
IV.1 IMPLEMENT INSTRUCTIONAL MANAGEMENT PLAN
IV.2 CONDUCT INSTRUCTION
V.1 CONDUCT INTERNAL EVALUATION
V.2 CONDUCT EXTERNAL EVALUATION
V.3 REVISE SYSTEM

Figure 2. 2nd Generation ISD Model (1970s).

© Branson (1976)

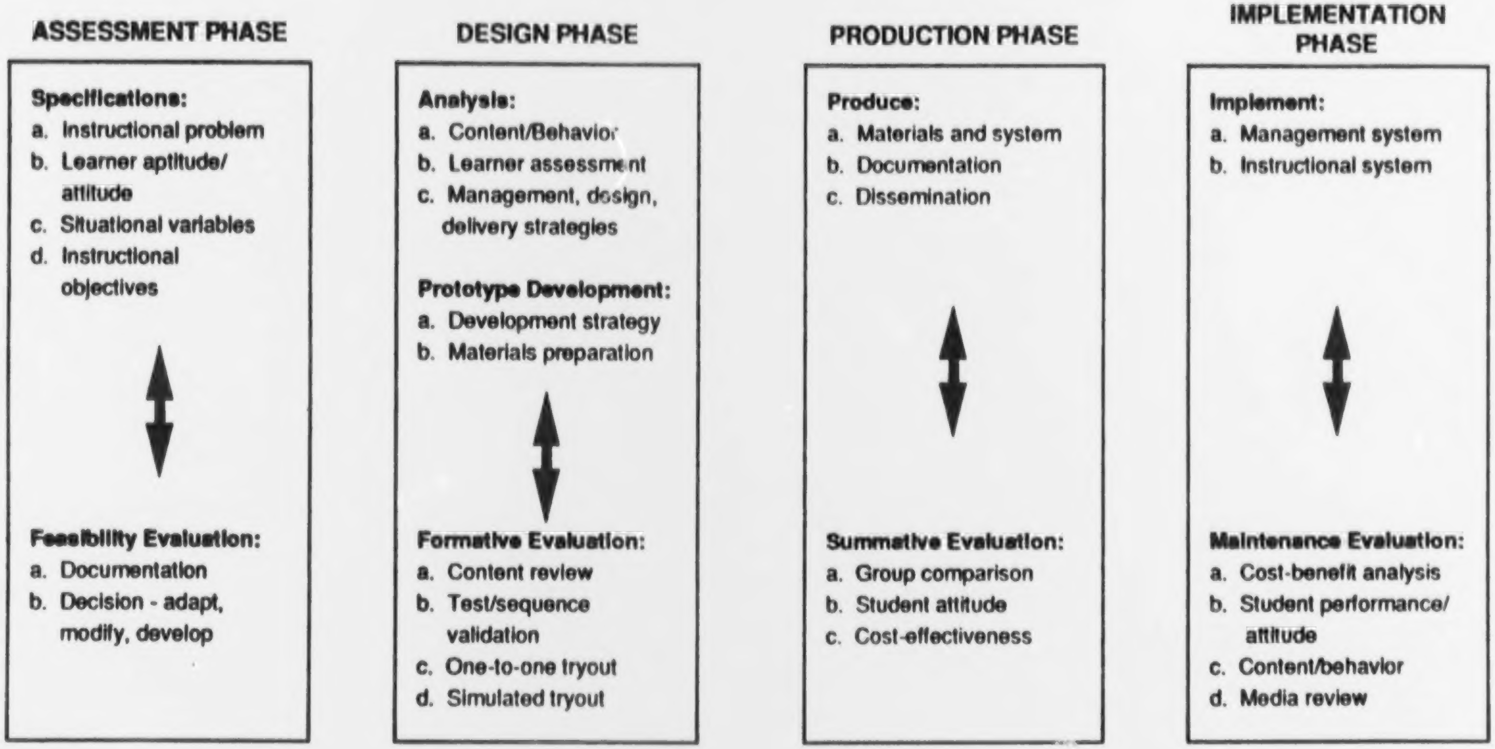
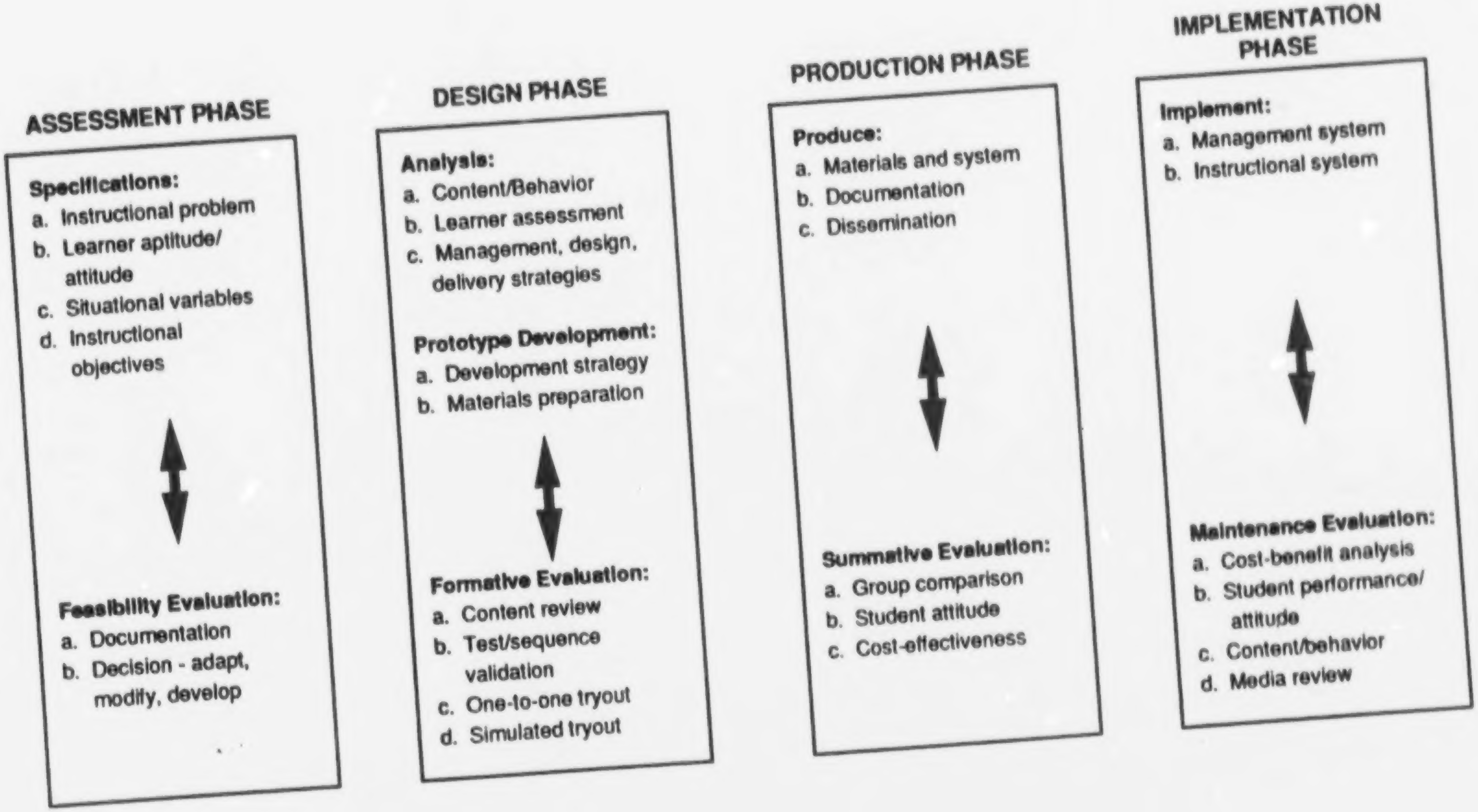

ASSESSMENT PHASE
Specifications:
a. Instructional problem
b. Learner aptitude/ attitude
c. Situational variables
d. Instructional objectives
Feasibility Evaluation:
a. Documentation
b. Decision - adapt, modify, develop
DESIGN PHASE
Analysis:
a. Content/Behavior
b. Learner assessment
c. Management, design, delivery strategies
Prototype Development:
a. Development strategy
b. Materials preparation
Formative Evaluation:
a. Content review
b. Test/sequence validation
c. One-to-one tryout
d. Simulated tryout
PRODUCTION PHASE
Produce:
a. Materials and system
b. Documentation
c. Dissemination
Summative Evaluation:
a. Group comparison
b. Student attitude
c. Cost-effectiveness
IMPLEMENTATION PHASE
Implement:
a. Management system
b. Instructional system
Maintenance Evaluation:
a. Cost-benefit analysis
b. Student performance/ attitude
c. Content/behavior
d. Media review

Figure 3. 3rd Generation ISD Model (1980s).

© 1990 Tennyson & Associates



ASSESSMENT PHASE

Specifications:
a. Instructional problem
b. Learner aptitude/ attitude
c. Situational variables
d. Instructional objectives

Feasibility Evaluation:
a. Documentation
b. Decision - adapt, modify, develop

DESIGN PHASE

Analysis:
a. Content/Behavior
b. Learner assessment
c. Management, design, delivery strategies

Prototype Development:
a. Development strategy
b. Materials preparation

Formative Evaluation:
a. Content review
b. Test/sequence validation
c. One-to-one tryout
d. Simulated tryout

PRODUCTION PHASE

Produce:
a. Materials and system
b. Documentation
c. Dissemination

Summative Evaluation:
a. Group comparison
b. Student attitude
c. Cost-effectiveness

IMPLEMENTATION PHASE

Implement:
a. Management system
b. Instructional system

Maintenance Evaluation:
a. Cost-benefit analysis
b. Student performance/ attitude
c. Content/behavior
d. Media review

Figure 3. 3rd Generation ISD Model (1980s).

© 1990 Tennyson & Associates

Title:

Two Theories of Design and Instructional Design

Author:

Steven D. Tripp

Two Theories of Design and Instructional Design

Steven D. Tripp
University of Kansas

Introduction

Nobel Laureate, Herbert Simon (1981) addressed the nature of fields like computer science, engineering, and education by proposing a difference between the natural sciences and what he called "sciences of the artificial." The four qualities that separate the natural sciences from the artificial or design sciences are: 1) artificial things are synthesized by man, 2) they imitate appearances of natural things but lack the reality of them, 3) artificial things can be characterized in terms of functions, goals and adaptation, and 4) artificial things are usually discussed in terms of imperatives as well as descriptives.

Simon wrote that, *"...the proper study of mankind is the science of design...."* (1969/1981, p. 159) Simon conceived of design as encompassing all kinds of human activity which involve planning. *"Engineers are not the only professional designers. Everyone designs who devises courses of action aimed at changing existing situations into preferred ones. The intellectual activity that produces material artifacts is no different fundamentally from the one that prescribes remedies for a sick patient or the one that devises a new sales plan for a company or a social welfare policy for a state. Design, so construed, is the core of all professional training; it is the principle mark that distinguishes the professions from the sciences. Schools of engineering, as well as schools of architecture, business, education, law, and medicine, are all centrally concerned with the process of design."* (1969/1981, p. 129)

Design is any purposeful making or doing that involves planning. Design encompasses most kinds of skilled behavior used in achieving goals under uncertain conditions. Teachers, administrators, politicians, doctors, engineers, lawyers and researchers are daily involved in design activities. Given the ubiquity of design-like situations it is surprising that so little is known about it.

The Problem.

What is the nature of design? How do skilled designers function? Can a theory of design be constructed which will allow novice and expert instructional designers to perform their tasks more efficiently and effectively? Although there are many opinions on these subjects, those many views can perhaps best be epitomized by two general theories of design.

1: Design as Optimization.

Simon conceived of design as an instance of problem-solving. Problem-solving involves moving the current situation progressively closer to the goal situation. A formal description of problem-solving involves representing the problem as a problem-space with initial, intermediate, and goal states. The solution to the problem involves searching for operators which will transform the initial state into the goal state. For ill-structured problems, heuristics rather than algorithms are required to achieve ends. A standard heuristic for the solving of difficult problems is means-ends analysis. Ends are defined and means to those ends are specified. If no means are apparent, the problem is decomposed into a hierarchy of sub-problems. This decomposition continues until means are discovered to solve the sub-problems. Thus, problem solving and therefore design, is a matter of finding the best description of the problem. In Simon's interpretation, a theory of design is equivalent to a formal representation of problem-solving heuristics.

2: Design as Dialogue

Donald Schön, a philosopher who has been involved with architects, has as a goal the development of an an epistemology of practice; that is, a characterization of the kinds of knowledge which a skilled designer/planner must possess. As a result of his observations of professional architects, he has proposed an alternative view of design. Schön (1988) has argued that the defining characteristics of design activities are *uncertainty, uniqueness* and *conflict* . Schön does not see design as heuristic problem-solving. On the contrary, design is a process of reflection-in-action; and designers take on the task of turning indeterminate situations into determinate ones (Schön, 1987).

Presented at the Annual Meeting of the AECT, Orlando, February 1991

Design, in this view, is a dialogue with phenomena. It can be seen as a kind of social process of negotiation. The objects of a design world are like Seymour Papert's objects to think with. We make worlds (in Goodman's sense) , Schön says, through a Gestalt-like process consisting of selective attention, grouping, boundary setting, and naming. By classifying we create **types**. Schön believes types are particulars that function in a general way. He asserts that the transaction between familiar type and unique design situation is a **metaphorical** process in which a designer both transforms a design situation and enriches the repertoire of types available to him for further design. Design, by this reading, has a hermeneutic quality. Problems do not have an objective existence, but rather are constructions created by designers as they grapple with uncertainty.

Implications

The main philosophical basis of modern instructional design has been the systems approach. The traditional interpretation of the systems approach has been much closer to Simon's view than Schön's. Thus instructional designers were trained in a means-ends analysis methodology and were expected to conform to this methodology in their professional life. Interestingly, Bela Banathy, an early advocate of the linear systems approach for instructional designers, has recently published a paper (Banathy, 1987) which aligns the systems approach more closely with the design-as-dialogue interpretation. He now sees design as holistic, interactive, spiralic, and dialectical. In spite of this change, the education of instructional designers still consists of exposure to hierarchical, top-down methods of problem decomposition. Additionally, empirical studies of novice and expert designers have indicated that in many cases, designers deviate from the hierarchical top-down type of design that Simon hypothesizes. Since both hierarchical and opportunistic approaches have both been observed in skilled designers, it may be that both Simon and Schön are, in a sense, correct, but they are describing different approaches used by people of differing abilities working on different kinds of problems.

Review of Literature

Empirical Studies of Design

There is a small, but growing, body of research on the activities of designers in action. This literature spans such areas as architecture, engineering, and computer science. The following table summarizes a portion of the published data on design outside of education.

Name & Date	Domain	Content
Adelson & Soloway (1984, 1985) in Guindon (1990)	Systems design	Three expert designers' work was systematic and balanced. Balanced means no part of the design was developed in significantly greater detail than other parts. Unbalanced design was followed only when a part was not familiar
Bucciarelli, L. L. (1988).	An ethnographic perspective on engineering design.	Studied design process within an engineering firm. Previous studies showed that a design team spent less than half of their time on legitimate design acts. Describes a discourse between three engineers and the artifact of design itself. Different aspects of an artifact are of interest to different persons. Differently schooled persons inhabit different "object worlds," each with its own symbols, metaphors, and models. Cites three cases of design discourse (specification, naming and decision making). Concludes: Values and beliefs do not derive from free choice; design change will not come by influencing young engineers. Ambiguity is always with us. Artifacts are not the design; they symbolize agreements. Knowledge is context-dependent but there will always be design moves that reach across boundaries.

Carroll & Rosson (1985)	General design	Examined empirical studies of designers in action and based on these studies they argue that the process is: (1) non-hierarchical, (2) neither strictly bottom-up nor top-down, (3) radically transformational, involving the development of partial and interim solutions which may ultimately play no role in the final design, and (4) intrinsically involving the discovery of new goals.
Cornforth (1976, in Lera, 1983)	Architectural design	Found subjects alternated between specification and search processes, contrary to standard methods. Designers develop a simple solution with a superficial plausibility and use that to develop a more detailed solution.
Curtis, B., Krasner, H., & Iscoe, N. (1988).	A field study of the software design process for large systems	Anecdotal report of the thinking of system designer on the design process. Software development is learning, negotiation, and communication process.
Darke (1979, in Lera, 1983)	Architectural design	Found that designers do not start with an explicit list of factors to be considered.
Downing, F. (1987).	Architecture	Studied architects working on a building project. Thematic, analogous and "charged" imagery creates a framework of thought. Prototypical solutions prestructure problems (design-by-precedent). The basic model of involves "conjecture-analysis." architects have an "image bank" of memorable solutions. Design is "seeing-as" (in Schon's terms).
Eastman (in Lera, 1983)	General design	Found a correspondence between representation used and constraints discovered. Instead of generating abstract attributes and relationships, subjects generated a design element and determined its qualities.
Foz (1972, in Lera, 1983)	Architectural design	Found designers use existing known examples to solve problems. Some manipulated 3-D models as if they were reality. Most skilled subject called on precedents, made many proposals, performed more tests, use more analogies deliberately, and was explicitly aware of his creative process. Whatever the designer is confident of being able to produce is put aside until it is needed. Designers deal with problems by consulting precedents.
Guindon, R, (1990)	Software design	Analysis of protocols of three software designers working on the Lift problem. Concludes that design process for expert designers is both opportunistic and top-down. Top-down is a special case when problem is well-structured. Makes recommendations for the design of environments that support opportunistic design.
Guindon, R., Curtis, B., & Krasner, H. (1987).	A model of cognitive processes in software design: An analysis of breakdowns in early design activities by individuals	Protocols of three experienced designers indicate three main sources of breakdowns: Lack of knowledge, cognitive limitations and combinations of those two factors. Designers exhibited huge individual differences in their design strategies and their design solutions. These finding could not be accounted for by prescriptive models.

Hykin (1972, in Lera, 1983)	Engineering design	Reports eleven case studies of engineers and concludes that it is impossible to isolate and identify simple design strategies. However he reported that exploration of alternatives led to a clearer understanding of the problem. Engineers expressed a need for a method of recording design decisions.
Jeffries, R., Turner, A. A., Polson, P. G., & Atwood, M. E. (1981).	The processes involved in designing software.	Studied four experienced software designers and five novices: Great variety of solutions both within and between levels. Novices lacked processes for solving subproblems and ways of representing knowledge effectively. Most experienced designer had many digressions. Some used "problem-solving-by-understanding."
Klein, G. A. (1987).	Various real-world design situations.	Reports several studies of designers dealing with difficult problems. Ill-defined problems require goal clarification and option development. Recognitional processes play a key role in design decision making as well as problem solving. Found little evidence of systematic use of decision analysis methods. Use of analogs lead to comparison-based prediction. Imagery is an important part of design. Research is used selectively and only to support preferences. Rapid prototyping is an attractive strategy for designers.
Lawson, B. (1980).	Architecture	Compared design strategies of final-year architectural students and science students at a comparable point in their education. Strategies were different but consistent. Scientists were problem-driven and sought to discover a rule. Architects were solution-driven and sought to learn about the problem from plausible solutions. In a second experiment with the same materials, high school and first year architecture students were tested. Both performed more poorly than the postgraduate students and neither group showed consistent patterns. Thus the consistent strategies of the architects were not a natural tendency.
Lee, T. Y., & Radcliffe, D. F. (1990).	Innate design abilities of first year engineering and industrial design students.	Two hundred twenty-six students studied by retrospective review. Experience improves design skills. Engineering and industrial design students exhibited clear differences of attitude, and this may be reflected in their choice of career. There appeared to be an "engineering way" of approaching design tasks and this tendency is acquired before entering university.
Lera, S. (1983).	Synopses of some recent published studies of the design process and designer behavior.	Lera (1980, in Lera 1983)) experimentally studied architects, architectural students, and non-architects and found that differences in design plans could be predicted by differences in subjective value weightings given before the design exercise.
Nadler, G. (1989).	Engineers and other professional designers.	Observational study of outstanding designers (engineer, architect, commercial artist, physician & lawyer) showed that they did NOT follow conventional methods, but rather a purpose and solution-after-next orientation.

Simmonds (1980, in Lera, 1983)	Architectural design	Studied twelve graduate students of architecture and found that they differed in their methods. Some analyzed the problem first, some generated solutions, and some looked at resources and constraints. Among those who analyzed the problem first, there was a variety of approaches. Some identified subproblems and attacked them in order of importance. Others generated a range of alternatives. A persistent problem was the inability to reverse the process of concretization. More successful students exhibited greater range and flexibility in their decisions.
Ullman, Stauffer, & Dietterich (1987) in Guindon (1990)	Mechanical engineering	Designers progress from systematic to opportunistic behaviors as design evolves.
Visser (1987) in Guindon (1990)	Software design	A team of programers showed opportunistic activities due to economic use of means, postponing decisions, handling familiar components, and changing decision criteria.

Summary

Design is highly varied and complex. Skilled designers may exhibit some common tendencies but they often do not follow normative models of design. Synthetic and analytic activities are often intertwined. Experience and precedent are often determining factors in design solutions. Many studies report opportunistic behavior.

Educational/instructional design

Instructional design as such has been little studied but there is a substantial literature on teacher planning. The classical model of curriculum planning is Tyler's. Essentially Tyler's model specifies that planning should specify goals, determine activities which address those goals, sequence the activities, and evaluate the results with respect to the goals. Empirical studies (Jackson, 1966,1968; Clark & Yinger, 1980; Zahoric, 1975) have generally found that teachers are not plan-driven but are reactive in class and that the teacher's hidden side is in the planning process. Teachers use intuitive planning and think in terms of doing. Teachers do not begin with objectives as Tyler would have them, but rather with activities. Taylor (1975) concluded teacher thinking is an inversion of theoretical thinking. Teachers first consider the context, then the situation, then the purposes.

Walker's (1971) summary of several curriculum development projects concluded similarly that objectives and even evaluation are not essential parts of curriculum development. Naturalistic views of projects indicated that objectives were a diversion or an appendix to the developers' work and not a starting point as the classical models suggest.

In an interesting use of ethno-methodology, Cain (1989) studied two preservice teachers. The two teachers were similar in background and beliefs. One was encouraged to use a "rational means-ends" planning model (based on Tyler). The other used a "Creative Planning Model" developed by Cain. The creative model consisted of three stages: preplanning, planning, and postplanning. The first stage involved observing, brainstorming, researching, and imagining a whole classroom environment. The second stage involved translating insights into goal, activities, and narratives. The last stage encouraged post hoc reflection and internalizing. The creative planner focused more on individual student characteristics, shared more teacher insights, and had more thoughts before teaching. The rational planner recalled more unplanned decisions. The creative planner reflected more after teaching. The rational planner thought of constraints as problems. The rational planner worked toward objectives even at the expense of her students. Cain suggests that the rational means-end model is not readily accepted by teachers and seems to stifle creative planning, therefore, other planning models may be desirable. Cain's data indicate that teaching results, at least, may be subject to the kinds of design methods utilized.

In virtually the *only* study in the literature on instructional design processes, Kerr (1981) studied teachers as instructional designers. Students in four summer courses worked on instructional designs and filled out questionnaires. Many reported thinking initially in terms of media rather than objectives. Students reported a high degree of uncertainty and a need to go back and reformulate objectives. Others reported that the materials suggested new possibilities. Kerr noted a general resemblance between the strategies of the teachers and the processes of artistic designers. Kerr concluded that instructional design models should encourage divergent thinking.

The only other known study of instructional designers is a dissertation by Wayne Nelson of Southern Illinois University. Nelson studied the early decision processes of four experienced instructional designers. Nelson found that while there was a discernable pattern in their methods, they did not follow Simon's top-down strategy.

As can be seen from the above, empirical studies of instructional design and teacher planning seem to indicate a strong resemblance between instructional design activities and other types of design. Instructional designers, like other designers, often do not use strongly hierarchical, top-down methods. Because of this resemblance, the results of studies of designers may have implications beyond their original domain.

Implications

Prescriptive design methodologies must support real-world methods in order to be effective. Real-world methods of instructional designers are not well understood, but if they resemble the methods of other kinds of designers, they are likely to opportunistic and non-uniform. In that case the tools we provide for instructional designers should support their preferred methods.

Guindon (1990) has argued that at least until the proper decomposition of the content is discovered the design process should be opportunistic. Only after the proper decomposition is known can a top-down approach be used. Instructional designers, like other designers, are expected to document their approach and this requires that some sort of systematic methods be reported.

Guidon has made recommendations for designers in a computer-environment:

1. The environment should not embody a method that locks designers into strict order of activities.
2. The environment should support rapid access and shifts between tools to represent and manipulate different kinds of objects.
3. It should support easy navigation between these objects.
4. The representation languages should support a smooth progression from informal to formal representations.
5. The environment should support easy editing and reorganization.
6. It should support the identification of the origin of requirements.
7. It should support the representation of interim or partial design objects.

The extent to which these recommendations apply to instructional design is unknown, but it is likely that requirements such as these will be relevant to any computer-based tools designed to support instructional design

Extended Bibliography

Akin, Ö. (1990). Necessary conditions for design expertise and creativity. **Design Studies, 11,** 107-113.

Andrews, D. H., & Goodson, L. A. (1980). A comparative analysis of models of instructional design. **Journal of Instructional Development, 3(4),** 2-16.

Asimow, M. (1962). **Introduction to design.** Englewood Cliffs, NJ: Prentice-Hall.

Ballay, J. M. (1987). An experimental view of the design process. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 65-82). New York: North-Holland.

Banathy, B. (1987). Instructional systems design. In R. Gagné (Ed.), **Instructional technology: Foundations** (pp. 85-112). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bergland, G. D. (1981). Structured design methodologies. In G. D. Bergland & R. D. Gordon (Eds.), **Tutorial on software design strategies** (pp. 297-315). Los Angeles, CA: IEEE Computer Society.

Boff, K. R. (1987). The tower of Babel revisited: On cross-disciplinary chokepoints in system design. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 83-96). New York: North-Holland.

Branson, R. K. (1975). **Interservice procedures for instructional systems development: Executive summary and model.** Tallahassee: Florida State University, Center for Educational Technology. (ED 122 022)

Briggs, L. J. (Ed.). (1977). **Instructional design.** Englewood Cliffs, NJ: Educational Technology Publications.

Broadbent, G. (1973). **Design in architecture.** London: John Wiley & Sons.

Bruns, G. R., & Gerhart, S. L. (1986). **Theories of design: An introduction to the literature (STP-068-86).** Austin, TX: MCC Software Technology Program.

Bucciarelli, L. L. (1988). An ethnographic perspective on engineering design. **Design Studies, 9,** 159-169.

Cain, B. N. (1989). With worldmaking, planning models matter. **English Education, 21,** 5-29.

Caplan, R. (1982). **By design.** New York: McGraw-Hill.

Card, S. K., Moran, T. P., & Newell, A. (1983). **The psychology of human-computer interaction.** Hillsdale, NJ: Lawrence Erlbaum.

Carroll, J. M. (1990). **The Nurnberg Funnel: Designing minimalist instruction for practical computer skill.** Cambridge, MA: MIT Press.

Carroll, J. M., & Kellogg, W. A. (1989). **Artifact as theory-nexus: Hermeneutics meets theory-based design.** In K. Bice & C. Lewis (Eds.), CHI '89 'Wings for the mind' (pp. 7-14). New York: Association for Computing Machinery.

Carroll, J. M., & Rosson, M. B. (1987). Paradox of the active user. In J. M. Carroll (Ed.), **Interfacing thought** (pp. 80-111). Cambridge: MIT Press.

Carroll, J. M., Thomas, J. C., & Malhotra, A. (1980). Presentation and representation in design problem-solving. **British Journal of Psychology, 71,** 143-153.

Carroll, J. M., Thomas, J. C., Miller, L. A., & Friedman, H. P. (1980). Aspects of solution structure in design problem solving. **American Journal of Psychology, 93,** 269-284.

Chan, C-S. (1990). Cognitive processes in architectural design problem solving. **Design Studies, 11,** 60-80.

Chatelier, P. R. (1987). Psychology or reality. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 97-111). New York: North-Holland.

Clark, C. M., & Yinger, R. J. (1980). **The hidden world of teaching: Implications of research on teacher planning (Research Series No. 77).** East Lansing, MI: The Institute for Research on Teaching.

Clark, R. E. (1989). The future of technology in educational psychology. In M. C. Wittrock & F. Farley (Eds.), **The future of educational technology** (pp. 91-106). Hillsdale, NJ: Lawrence Erlbaum Associates.

Conklin, J. (1986). **A theory and tool for coordination of design conversations (STP-236-86).** Austin, TX: MCC Software Technology Program.

Cooper, R. S. (1987). Effects of technological and organizational trends on system design. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 19-41). New York: North-Holland.

Crinnion, J. (1989). A role for prototyping in information systems design methodology. **Design Studies, 10,** 144-150.

Cross, N. (1990). The nature and nurture of design ability. **Design Studies, 11,** 127-140.

Cross, N. (Ed.). (1984). **Developments in design methodology.** Chichester: John Wiley & Sons.

Curtis, B., Krasner, H., & Iscoe, N. (1988). **A field study of the software design process for large systems (STP-233-88).** Austin, TX: MCC Software Technology Program.

Dick, W., & Carey, L. (1985). **The systematic design of instruction (2nd ed.).** Glenview, IL: Scott, Foresman & Co.

Dinham, S. M. (1989). Teaching as design: Theory, research, and implications for design teaching. **Design Studies, 10,** 80-88.

Downing, F. (1987). Imagery and the structure of design inquiry. **Journal of Psycholinguistic Research, 11,** 61-86.

Dreyfuss, H. (1974). **Designing for people.** New York: Grossman Publishers.

Eder, W. E. (Ed.). (1987). **Proceedings of the 1987 international conference on engineering design (Vol. 1).** New York: The American Society of Mechanical Engineers.

Eder, W. E. (Ed.). (1987). **Proceedings of the 1987 international conference on engineering design (Vol. 2).** New York: The American Society of Mechanical Engineers.

Faiola, T., & DeBloois, M. L. (1988). Designing a visual factors-based screen display interface: The new role of the graphic technologist. **Educational Technology,** August.

Ferguson, D. L., & Henderson, P. B. (1987). The design of algorithms. **Machine-Mediated Learning, 2,** 67-82.

Floyd, C., Mehl, W-M., Reisin, F-M., Schmidt, G., & Wolf, G. (1989). Out of Scandinavia: Alternative approaches to software design and system development. **Human-Computer Interaction, 4,** 253-350.

Freeman, P. (1983). Fundamentals of design. In P. Freeman & A. I. Wasserman (Eds.), **Tutorial on software design techniques** (pp. 2-22). Los Angeles, CA: IEEE Computer Society.

Gagné, R. M., & Briggs, L. J. (1979). **Principles of instructional design (2nd ed.).** New York: Holt, Rinehart and Winston.

Gagné, R. M., & Merrill, M. D. (1990). Integrative goals for instructional design. **Educational Technology Research and Development, 38,** 22-30.

Glaser, R. (1976). Components of a psychology of instruction: Toward a science of design. Review of **Educational Research, 46,** 1-24.

Glynn, S. (1985). Science and perception as design. **Design Studies, 6,** 122-126.

Grant, D. P. (1977). How to operate an issue-based information system (IBIS) in support of a research and development project. **Design Methods and Theories, 11,** 221-237.

Grant, D. P. (1977). How to use the IBIS procedure for deliberation and argument in environmental design and planning. **Design Methods and Theories, 11,** 185-220.

Guindon, R. (1988). **A framework for building software development environments: System design as ill-structured problems and as an opportunistic process (STP-298-88).** Austin, TX: MCC Software Technology Program.

Guindon, R. (1990). Designing the design process: Exploiting opportunistic thoughts. **Human-Computer Interaction, 5,** 305-344.

Guindon, R., Curtis, B., & Krasner, H. (1987). **A model of cognitive processes in software design: An analysis of breakdowns in early design activities by individuals (STP-283-87).** Austin, TX: MCC Software Technology Program.

Hill, P. H. (1970). **The science of engineering design.** New York: Holt, Rinehart and Winston.

Hooper, S., & Hannafin, M. J. (1988). Learning the ROPES of instructional design: Guidelines for emerging interactive technologies. **Educational Technology, July,** 14-17.

Hubka, V. (1982). **Principles of engineering design** (W. E. Eder Trans.). London: Butterworth Scientific. (Original work published in 1980)

Hubka, V. (1983). Design tactics = methods + working principles for design engineers . **Design Studies, 4,** 188-195.

Hunt, R. M. (1987). The difficulties of design problem formulation. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 145-157). New York: North-Holland.

Ingram, A. L. (1988). Instructional design for heuristic-based problem solving. **Educational Communication and Technology Journal, 36,** 211-230.

Jackson, P. (1966). **The way teaching is.** Washington: National Educational Association.

Jackson, P. (1968). **Life in classrooms.** New York: Holt.

Jeffries, R., Turner, A. A., Polson, P. G., & Atwood, M. E. (1981). The processes involved in designing software. In J. R. Anderson (Ed.), **Cognitive skills and their acquisition** (pp. 255-283). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jonassen, D. H., & Hannum, W. H. (1987). Research-based principles for designing computer software. **Educational Technology,** December, 7-14.

Jones, J, C. (1984). **Essays in design.** Chichester, UK: John Wiley & Sons.

Jones, J. C. (1963). A method of systematic design. In J. C. Jones & D. G. Thornley (Eds.), **Conference on design methods** (pp. 53-73). Oxford: Pergamon.

Kendall, S. (1989). On design reasoning: Lessons from teaching architectural technology. **Design Studies, 10,** 89-95.

Kerr, S. T. (1981). How teachers design their materials: Implications for instructional design. **Instructional Science, 10,** 363-378.

Klein, G. A. (1987). Analytical versus recognitional approaches to design decision making. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 175-186). New York: North-Holland.

Klein, G. A. (1987). Applications of analogical reasoning. **Metaphor and Symbolic Activity, 2,** 201-218.

Knirk, F. G., & Gustafson, K. L. (1986). **Instructional technology.** New York: Holt, Rinehart and Winston.

Lawson, B. (1980). **How designers think.** Westfield, NJ: Eastview Editions.

Leifer, L. (1987). On nature of design and environment for design. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 211-220). New York: North-Holland.

Lera, S. (1983). Synopses of some recent published studies of the design process and designer behavior. **Design Studies, 4,** 133-140.

Levy, R. (1985). Science, technology and design. **Design Studies, 6,** 66-72.

Logan, B. S. (1989). Conceptualizing design knowledge. **Design Studies, 10,** 188-195.

MacLean, A., Young, R. M., & Moran, T. P. (1989). Design rationale: The argument behind the artifact. In K. Bice & C. Lewis (Eds.), **CHI '89 'Wings for the mind'** (pp. 247-252). New York: Association for Computing Machinery.

Maher, J. H., & Ingram, A. L. (1989, February). **Software engineering and ISD: Similarities, complementarities, and lessons to share.** Paper presented at the annual meeting of Association for Educational Communications and Technology, Dallas, TX.

Marca, D. A., & McGowan, C. L. (1988). **SADT ™ Structured analysis and design technique.** New York: McGraw-Hill.

Martin, E. A. (1987). Toward a more systematic, efficient design process: The potential impact of intelligent design aids. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 221-227). New York: North-Holland.

Mayall, W. H. (1979). **Principles in design.** New York: Van Nostrand Reinhold.

Meister, D. (1987). A cognitive theory of design and requirements for a behavioral design aid. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 229-244). New York: North-Holland.

Merrill, D. M. (1987, April). **Instructional design expert.** Paper presented at the annual meeting of American Educational Research Association, Washington, DC.

Merrill, M. D. (1987). The new component design theory: Instructional design for courseware authoring. **Instructional Science, 16,** 19-34.

Merrill, M. D., Li, Z., & Jones, M. K. (1989, November). **ID2: Second generation instructional design.** Paper presented at the annual meeting of ADCIS, Washington, DC.

Nadler, G. (1989). Design processes and their results. **Design Studies, 10,** 124-127.

Nunan, T. (1983). **Countering educational design.** New York: Nichols Publishing.

Oxman, R. (1990). Prior knowledge in design: A dynamic knowledge-based model of design and creativity. **Design Studies, 11,** 31-38.

Papanek, V. (1988). The future isn't what it used to be. **Design Issues,** 5, 4-17.

Pask, G. (1963). The conception of a shape and the evolution of a design. In J. C. Jones & D. G. Thornley (Eds.), **Conference on design methods** (pp. 153-167). Oxford: Pergamon.

Perkins, D. N. (1986). **Knowledge as design.** Hillsdale, NJ: Lawrence Erlbaum.

Peters, L. J., & Tripp, L. L. (1981). Comparing software design methodologies. In G. D. Bergland & R. D. Gordon (Eds.), **Tutorial on software design strategies** (pp. 316-321). Los Angeles, CA: IEEE Computer Society.

Petroski, H. (1989). Failure as a unifying theme in design. **Design Studies, 10,** 214-218.

Porter, W. L. (1988). Notes on the inner logic of designing: Two thought-experiments. **Design Studies, 9,** 169-180.

Potts, C., & Bruns, G. (1987). **Recording the reasons for design decisions (STP-304-87).** Austin, TX: MCC Software Technology Program.

Pugh, S. (1989). Knowledge-based systems in design activity. **Design Studies, 10,** 219-227.

Radcliffe, D. F., & Lee, T. Y. (1989). Design methods used by undergraduate engineering students. **Design Studies, 10,** 199-207.

Reigeluth, C. M. (1983). **Instructional-design theories and models.** Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M. (1987). **Instructional-theories in action.** Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M. (1989). Educational technology at the crossroads: New mindsets and new directions. **Educational Technology Research and Development, 37,** 67-80.

Reigeluth, C. M., & Darwazeh, A. (1982). The elaboration theory's procedure for designing instruction. **Journal of Instructional Development, 5,** 22-31.

Reigeluth, C. M., & Stein, F. S. (1983). The elaboration theory of instruction. In C. M. Reigeluth (Ed.), **Instructional-design theories and models** (pp. 335-381). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reigeluth, C. M., Bunderson, C. V., & Merrill, M. D. (1978). What is the design science of instruction? **Journal of Instructional Development, 1,** 11-16.

Reisner, P. (1981). Formal grammar and human factors design of interactive graphics system. **IEEE Transactions on Software Engineering, SE-7,** 229-240.

Richards, J. N. J., Bez, H. E., Gittins, D. T., & Cooke, D. J. (1986). On methods for interface specification and design. **International Journal of Man-Machine Studies, 24,** 545-658.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. **Policy Sciences, 4,** 155-169.

Rittel, H. W. J., & Webber, M. M. (1974). Wicked problems. In N. Cross, D. Elliott, & R. Roy (Eds.), **Man-made futures** (pp. 272-280). London: Hutchinson & Co..

Romiszowski, A. J. (1981). **Designing instructional systems.** New York: Nichols Publishing.

Rouse, W. B. (1987). Designers, decision making, and decision support. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 275-283). New York: North-Holland.

Rouse, W. B., & Boff, K. R. (1987). Designers, tools, and environments: State of knowledge, unresolved issues, and potential directions. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 43-63). New York: North-Holland.

Rowe, P. G. (1987). **Design thinking.** Cambridge: MIT.

Rzevski, G. (1984). Prototypes versus pilot systems: Strategies for evolutionary information system development. In R. Budde, K. Kuhlenkamp, L. Mathiassen, & H. Züllighoven (Eds.), **Approaches to prototyping** (pp. 356-367). Berlin: Springer-Verlag.

Sage, A. P. (1987). Knowledge, skills, and information requirements for systems design. In W. B. Rouse & K. R. Boff (Eds.), **System design** (pp. 285-303). New York: North-Holland.

Salisbury, D. F. (1990). General systems theory and instructional systems design. **Performance & Instruction, 29(2),** 1-11.

Schneiderman, B. (1987). **Designing the user interface.** Reading, MA: Addison-Wesley.

Schön, D. A. (1987). **Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.** San Francisco: Jossey-Bass.

Schön, D. A. (1988). Designing: Rules, types and worlds. **Design Studies, 9,** 181-190.

Simon, H. A. (1981). **The sciences of the artificial (2nd ed.).** Cambridge, MA: MIT Press. (Original work published in 1969.)

Spillers, W. R. (Ed.). (1974). **Basic questions of design theory.** New York: American Elsevier.

Stauffer, L. A., & Ullman, D. G. (1988). A comparison of the results of empirical studies into the mechanical design process. **Design Studies, 9,** 107-114.

Sternberg, R. J. (1990). **Thinking styles: Keys to understanding student performance.** Phi Delta Kappan, 71, 366-371.

Swanson, E. B. (1977). A methodology for IBIS data gathering and development. **Design Methods and Theories, 11,** 256-261.

Ullmer, E. J. (1988). On the form of human enterprise: Examining the design of design. **Educational Technology, September,** 14-21.

Walker, D. (1971). A naturalistic model for curriculum development. **School Review, 80,** 51-65.

Walters, R. J. (1985). Informed, well-ordered and reflective: Design inquiry as action research. **Design Studies, 7,** 2-13.

Whitefield, A., & Warren, C. (1989). A blackboard framework for modelling designers' behaviour. **Design Studies, 10,** 179-187.

Whiteside, J., & Wixon, D. (1987). Discussion: Improving human-computer interaction -a quest for cognitive science. In J. M. Carroll (Ed.), **Interfacing thought** (pp. 353-365). Cambridge: MIT Press.

Whitten, J. L., Bentley, L. D., & Barlow, V. M. (1989). **Systems analysis & design methods (2nd ed.).** Homewood, IL: Irwin.

Willem, R. A. (1988). On knowing design. **Design Studies, 9,** 223-228.

Wilson, I. G., & Wilson, M. E. (1965). **Information, computers and system design.** New York: John Wiley & Sons.

Wixon, D., Holtzblatt, K., & Knox, S. (1990). Contextual design: An emergent view of system design. In J. C. Chew & J. Whiteside (Eds.), **Empowering People CHI '90 Conference Proceedings** (pp. 329-335). New York: Association for Computing Machinery.

Zahoric, J. (1975). Teacher planning models. **Educational Leadership, 33,** 134-139.

Title:

Color realism and Hemispheric Laterality

Authors:

Esther Waltz
Louis H. Berry

COLOR REALISM AND HEMISPHERIC LATERALITY

Esther Waltz
Community College of Allegheny County
Pittsburgh, PA

Louis H. Berry
University of Pittsburgh
Pittsburgh, PA

Extensive research relative to the use of color as a variable in enhancing learning has led to theories, albeit conflicting about the processing of color information. Other research has utilized findings pertinent to cerebral laterality to suggest how color is stored in memory. Little research, however, has focused on the interaction between color memory and hemispheric lateralization.

The purpose of this study was, therefore, to investigate the interaction between cognitive lateral functions and pictorial recognition memory for pictures presented in three different color modes: realistic color, non-realistic color and monochrome (black and white).

BACKGROUND RESEARCH

Substantial research has focused on the role of color in visual learning (Dwyer, 1972, 1987; Chute, 1979; Lamberski, 1980). Research reported by Berry & Dwyer (1982) and Berry (1982) investigated the relative effectiveness of realistic and non-realistic color in visual learning tasks. Findings showed that in different tasks, either realistic or non-realistic visuals were superior to monochrome materials These studies suggest that when color is used in a realistic manner, it can supply additional, meaningful verbal or linguistic labels, while the use of color simply as a coding technique (non-realistic) may serve only to increase the overall number of cues, but not provide additional realistic attributes or associations with which to store or retrieve information.

Other researchers have explored the degree of processing of visual information in the right and left cerebral hemispheres (Hellige, 1980, 1983; Bryden, 1982: Corballis, 1983). This research has suggested that tasks may be prioritized to particular locations in either the left or right hemispheres; the left hemisphere associated with verbal processes and the right with visuospatial processes (Bogen, 1977; Gazzaniga, 1985; Sperry, 1984). Hemispheric specialization became a topic of interest after Sperry's conclusions that two separate sides were responsible for distinct cognitive processes (Levy, 1983; Springer & Deutsch, 1985). Studies concerned with abnormal subjects concluded that functions of the right hemisphere were critical for color discrimination (DeRenzi & Spinnler, 1967; Nikolaenko, 1982; Scotti & Spinnler, 1970). Whenever normal subjects' visual fields were tested, the data suggested that some degree of independence of visual analysis and color recognition existed in each half of the brain (Dimond & Beumont 1972). Davidoff (1976) reported a right hemisphere advantage in perception of color stimuli. Questions regarding the lateral functions of the brain were further illuminated in the results of Chute's (1980) study which separated subjects into high visuospatial learners and low visuospatial learners. Color inhibited learning for the low-spatial group and increased learning for the high-spatial group. The implication was that those who were classified as having a low-spatial aptitude might profit less from color information than those classified as being visual or of high-spatial aptitude. Beauvois and Salillant (1985) suggested that color was represented in two separate memory systems; one classified as verbal and the other as visual. This suggestion supported the classic dual coding theory of Paivio (1971) that posits two intradependent but functionally independent systems for learning.

Evidence suggests that many lateral functions tend to cluster in hemispheres with the left specialized for syntactic and verbal skills and the right for visuospatial activities (Gordon, 1986; Hiscock & Kinsbourne, 1987). The left visual half-field/right hemisphere advantage is usually found for visuospatial tasks (Hellige, 1980). However, individuals are lateralized for speech whether right or left-handed with the majority being left hemisphere specialized.

It appears that the perception of color is basically a right hemisphere function although this has not been totally resolved. Results of color studies in which researchers utilized the left and right visual fields to study functions of the hemispheres are at times conflicting. Processing visual information conveyed by color can be affected by a number of variables e.g. both realistic and nonrealistic cues, verbal sequential and visuospatial interrelationships, and task complexity. When tasks are at a low level of difficulty, bilateral activation may be at a low level; whereas as task complexity increases, levels of cognitive processing and integration of hemispheric functions may be enhanced. Although evidence clearly disputes that individuals learn with only one side of the brain, there is some evidence of individual differences relative to hemispheric functions. In addition, many experimental methods have yielded insights into human cognition in which the brain can be viewed as having separate but interacting parts (Gazzaniga, 1989).

Berry (1990) suggested that the role of color in visuals could function in a dual role as either a cue to provide additional information without any realistic attributes or it could serve to convey those attributes of realism associated with the information.

METHOD

The study utilized as its basis, stimulus materials developed and used by Berry (1982, 1990). These were slides selected from a pool of travel and general geographic scenery slides taken in various parts of the United States and Canada. Slides with any verbal material, unique objects, or recognizable human figures were excluded. The entire collection of materials was randomly divided into thirds. One group was retained in he original realistic color format and a second group was produced in nonrealistic color by means of photographic reversal. In this way, the total number of visual cues were held constant while the degree of realism varied. In addition, a third group was recopied onto black and white slides. The original slides were randomly assigned to be copied in such a manner that a total of sixty split-slides were produced. The sixty split-sides allowed for simultaneous projection of two scenes with one of the left visual field and the other in the right visual field. To insure that each image was presented to the proper visual field, the paired images were separated on the slide a distance equal to four degrees of radius at the projection distance. This resulted in a space of 3.6 inches between each image as viewed by the subject. In this way, the right and left visual fields corresponded to the image size and did not overlap.

A sample of thirty, right-handed, male subjects were drawn from the student population of a Pennsylvania community college. All subjects were volunteers.

The list learning procedure was employed to present the stimulus materials. Subjects were seated approximately thirty inches from an 18.75 in. x 12.50 in. rear projection screen. A Kodak Ektagraphic Slide projector was used to rear project the stimulus slides using an Ilex No. 4 Synchro Electronic shutter to control the amount of time the slides were projected onto the screen. The sixty paired treatment slides were shown to subjects for a period of 200 ms each to eliminate saccadic eye movements. Each individual was instructed to concentrate on a dot at the center of the screen to ensure that the right and left visual fields were properly positioned to observe the respective images. Following the treatment presentation, subjects were presented a random distribution of all individual images (stimulus and distractor) for 5 sec each and asked to respond by saying "new" if the image was seen for the first time or "old" if it had been seen before. Verbal responses were recorded by the researcher.

The design of the study was a two-way repeated measures with two levels (left and right) of the visual field/hemisphere factor and three levels (realistic, nonrealistic, and monochrome) of the color treatment.

ANALYSIS

Analysis of the data obtained was conducted via the method of Signal Detection Analysis (Swets, 1964). This method has been advocated as a reliable means of assessing pictorial recognition memory (Berry, 1982, 1990; Loftus & Kallman, 1979).

The mean probability of hits and the mean error rates for each treatment and level were calculated as well as the measure of memory sensitivity, d', which was determined from tables developed by Elliot (1964). The d'statistic was transformed to eliminate negative values. These data are presented in Table 1.

Insert Table 1 about here

Test reliability coefficients were computed from standard errors and found to be 0.95 or greater for all variables. Analysis of variance procedures for two factor repeated measures were conducted on the d'adjusted, probability of hits, and total error scores. With respect to the analysis of the d'adjusted values, a significant F value (F=14.90, p<.001) for the interaction of Visual Field and Color was obtained as well as a significant main effect for color (F=3.85, p<.03). Analysis of the cell means via the Scheffé method indicated that a significant difference existed between the realistic and black and white treatments in the left visual field. This suggested the superiority of nonrealistic color in the left visual field/right hemisphere.

For the probability of hits, a significant interaction (F=14.16, p<.0001) between type of color and visual field was obtained and a significant F for the color main effect (F=3.55, p<.034). A Scheffé analysis confirmed that the nonrealistic color treatment was significantly higher than the black and white treatment in the left visual field. This finding would suggest that nonrealistic color materials were recognized better than black and white materials in the left visual field/right hemisphere.

Analysis of variance on the error rates produced a significant F value for the interaction (F=13.77, p<.001) and a significant F value for the main effect of color (F=3.85, p<.02). *Post hoc* comparisons via the Scheffé method indicated that the realistic color materials produced significantly more errors than did the nonrealistic color materials in the left visual field/right hemisphere. The nature of the three interactions are illustrated in Figures 1,2, & 3.

Insert Figures 1, 2, & 3 about here

In interpreting the findings of this study, it should be remembered that each visual field projects to the contralateral hemisphere of the brain. Consequently, the performance related to recognition in the left visual hemifield is processed in the right cerebral hemisphere and right visual hemifield performance is associated with left hemisphere processing.

The major findings of this study indicated that:

1. Significant interactions occurred between visual fields and modes of color.
2. There was no significant difference for memory of realistic color in either visual field.
3. There was no significant difference for black and white memory in either visual field.
4. Nonrealistic color was remembered significantly better than black and white in the left visual field.
5. There were fewer errors relative to processing of nonrealistic color in the right hemisphere than there were associated with realistic color.

The implications of these findings are that both realistic color and black and white processing are more functions of the left than the right hemisphere while nonrealistic color processing appears to be more a function of the right than the left hemisphere.

DISCUSSION

The main issue addressed in this study was the relationship between hemispheric lateralization and color mode. Nonrealistic color appears to be processed primarily in the right hemisphere rather than in the left and may serve merely as a cuing device. These findings are supported by the work of Jorgenson, Davis, Opella & Angerstein (1980) which found that color and word responses are left hemisphere processes while color alone involves additional right hemisphere processing. If, indeed, the functions of the right and left hemisphere are divided into two separate systems, one visual and one verbal, where visual is primarily a right hemisphere function and verbal a left hemisphere function, then the processing of nonrealistic color is more a function of the right hemisphere than the left hemisphere. It would appear, therefore, that those individuals classified as being primarily visuospatial (Chute, 1980) would recall color information, whether nonrealistic or realistic, better than monochrome. Realistic color may provide additional verbal components making it more a complex function of the left hemisphere as well as the right. These conclusions are supported by (Gazzaniga, 1989), who suggested that, as information becomes more complex, both intrahemispheric and interhemispheric functions become more unified. These conclusions are also supported by the recent findings of Berry (1990) in which subjects were given a complex verbal masking task in order to inhibit the processing of visual information in the left hemisphere. The patterns of interactions for d'adjusted responses, hit responses, and error scores were highly similar to those obtained in this study.

The findings suggest that nonrealistic color facilitates recognition more so in the right hemisphere than in the left hemisphere, while realistic and black and white visuals are less effective in promoting recognition in the right hemisphere. Nonrealistic color when presented in visuals is primarily a cuing device, while realistic color and black and white visuals have both visual and verbal components any may require higher levels of processing. It appears that as the components of visuals increase in complexity, lateral functions become less specialized and more unified.

REFERENCES

Beauvois, M. & Saillant, B. (1984). Optic aphasia for colours and colour agnosia: a distinction between visual and visuoverbal impairments in the processing of colours. Cognitive Neuropsychology, 2, 1-48.

Berry, L. H. (1990). Effects of hemispheric laterality on color information processing. Perceptual and Motor Skills, 71, 987-993.

Berry, L. H. (1982). Signal detection analysis of color and realism data. Reprint from Instructional Communications and Technology Research Reports, 13(1), n.p.

Berry, L. H. & Dwyer, F. M. (1982). Interactive effects of color realism and learners' IQ on effectiveness of visual instruction. Perceptual and Motor Skills, 54, 1087-1091.

Bogen, J. E. (1977). Some educational aspects of hemispheric specialization. In M. C. Wittrocks (Ed) The Human Brain. New Jersey: Prentice-Hall.

Bryden, M. P. (1982). Laterality: Functional asymmetry in the intact brain . New York: Academic Press.

Chute, A. G. (1979). Analysis of the instructional functions of color and monochrome cueing in media presentations. Educational Communications and Technology Journal. 27, 251-263.

Chute, A. G. (1980). AECT/RTD Young research paper: Effect of color and monochrome versions of a film on incidental and task-relevant learning. Educational Communications and Technology Journal, 28(1), 10-18.

Corballis, M. C. (1983). Human laterality . New York: Academic Press.

Davidoff, J. (1976). Hemispheric sensitivity differences in the perception of colour. Quarterly Journal of Experimental Psychology, 28, 387-394.

De Renzi, E. & Spinnler, H. (1967). Impaired performance on color tasks in patients with hemisphere damage. Cortex, 3, 194-217.

Dimond, S. & Beaumont, G. (1972). Hemisphere function and color naming. Journal of Experimental Psychology, 96, 87-91.

Elliot, P. B. (1964). Tables of d' in J. A. Swets, (Ed), Signal detection and recognition by human observers. New York: John Wiley & Sons.

Gazzaniga, M. S. (1989). Organization of the human brain. Science, 245, 947-952.

Gordon, H. W. (1986). The cognitive laterality battery: Tests of specialized cognitive functions. International Journal of Neuroscience, 29, 223-244.

Hellige, J. B. (1980). Cerebral hemisphere asymmetry: Methods, issues, and implications. ECTJ, 28, 83-98.

Hellige, J. B. (Ed.) . (1983). Cerebral hemisphere asymmetry: Method, theory and
application. New York: Praeger.

Hiscock, M., & Kinsbourne, M. (1987). Specialization of the cerebral hemispheres: Implications for learning. Journal of Learning Disabilities, 20, 130-143.

Jorgenson, C., Davis, J., Opella, J., & Angerstein, G. (1980). Hemispheric asymmetry in the processing of Stroop stimuli: An examination of gender, hand-preference, and language differences. International Journal of Neuroscience, 11, 165-169.

Lamberski, R. J. (1980). A comprehensive and critical review of the methodology and findings in color investigations. Paper presented at the annual convention of the Association for Educational Communications and Technology, Denver, CO.

Levy, J. (1983). Research synthesis on right and left hemisphere: We think with both sides of the brain. Educational Leadership, 95, 61-78.

Loftus, G. R., & Kallman, H. J. (1979). Encoding and use of detailed information in picture recognition. Journal of Experimental Psychology: Human Learning and Memory, 5, 197-211.

Nikolaenko, M. M. (1982). Role of the dominant and non-dominant hemispheres in the perception and naming of color, Human Physiology, 7(3), 441-448.

Paivio, A. (1971). A dual coding approach to perception and cognition. In Pick and Saltzman, Modes of Perceiving and Processing Information. New York: Wiley.

Scotti, G., & Spinnler, H. (1970). Color imperception in unilateral hemisphere-damaged patients. Journal of Neurology, Neurosurgery, and Psychiatry, 33, 22-28.

Sperry, R. (1984). Consciousness, personal identity and the divided brain. Neuropsychologia, 22(6), 661-673.

Springer, S., & Deutsch, G. (1985). Left brain, right brain. New York: W. H. Freeman.

Swets, J. A. (Ed.) (1964). Signal detection and recognition by human observers. New York: Wiley.

Table 1.
Means and Standard Deviations for Probability of Hits, Error Rates and d' Adjusted, by Treatments and Visual Field/Hemispheric Localization. (N = 30)

Treatment	P(H) M	P(H) S.D.	Errors M	Errors S.D.	d' Adjusted M	d' Adjusted S.D.
Left Visual Field / Right Hemisphere						
Realistic Color	.512	.126	16.22	2.10	2.03	.379
Nonrealistic Color	.627	.200	13.33	2.96	2.31	.468
Black & White	.432	.190	15.28	2.84	1.80	.487
Right Visual Field / Left Hemisphere						
Realistic Color	.580	.121	14.95	2.30	2.21	.415
Nonrealistic Color	.529	.186	15.27	2.30	2.06	.356
Black & White	.508	.189	13.78	2.86	2.05	.476

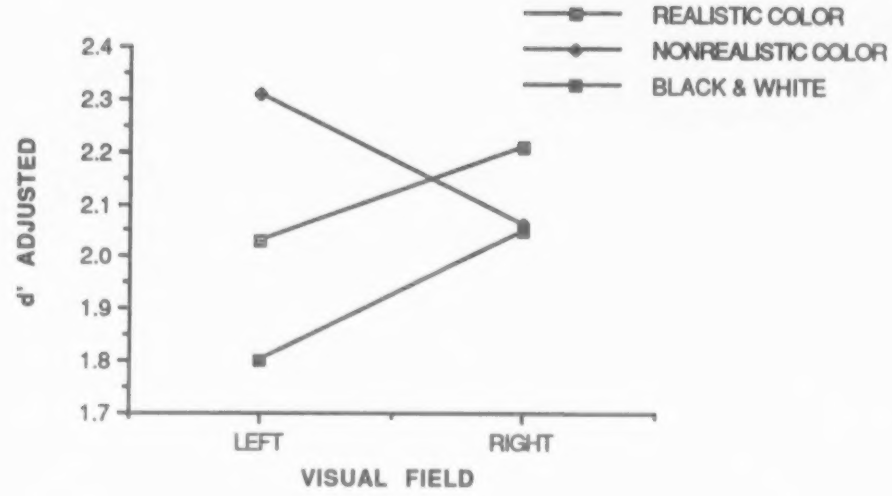


Figure 1. Mean d' Adjusted values for color treatments (realistic, nonrealistic, black & white) by visual fields.

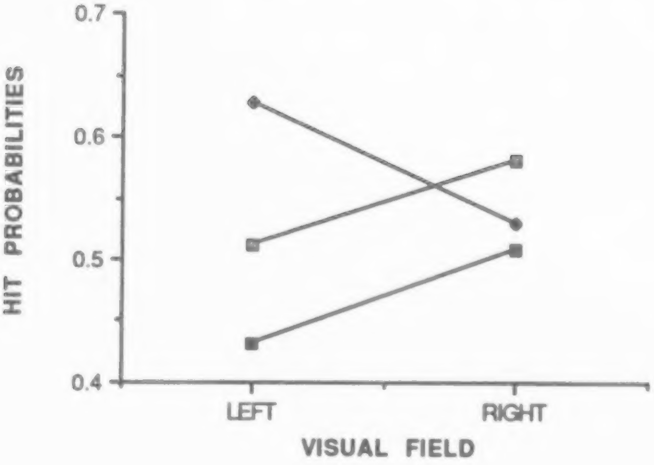


Figure 2. Mean hit probabilities for color treatments (realistic, nonrealistic, black & white) by visual fields.

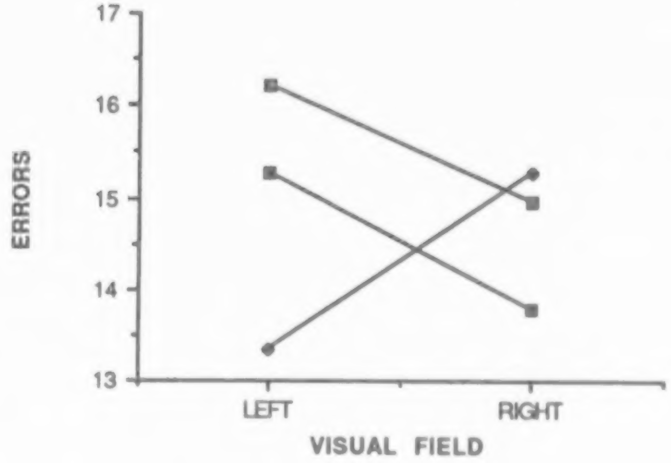


Figure 3. Mean error values for color treatments (realistic,nonrealistic, black & white) by visual field.

Title:

Instructor's Plan: A Lesson Planning Expert System for School Teachers

Authors:

Dennis A. Wilkins
Paul F. Cook
Edward E. Green

Instructor's Plan (IP) is an expert system for rapid lesson design and authoring. It is the product of several years of research and development at Brigham Young University and has received about two years of beta-testing by a large number of instructors (Wilkins, 1990). Its primary audience is preservice and inservice school teachers, with potential secondary audiences of higher education professors and corporate trainers. The main purpose of IP is to provide instructors with an intelligent productivity tool and job aid. A secondary purpose is to teach instructors good design principles. IP will run on IBM PC, XT, AT, and PS/2 systems and compatibles with a color monitor and at least one floppy disk drive.

The central feature of IP is its expert system possessing the primary components typically found in other expert systems, namely the knowledge base and inference engine (Townsend & Feucht, 1986). IP's knowledge base has two main aspects contained in several files. First, the strategy selection aspect of the knowledge base resides in an executable file along with the rest of the compiled source code. Here knowledge is represented in the form of production rules (Bahrami, 1988). Many of these rules contain certainty factors used in the inferencing process of the expert system (Levine, Drang, & Edelson, 1988). This allows IP to deal with a considerable amount of uncertainty related to the instructional conditions at hand. Second, the instructional strategies aspect of the knowledge base is contained in structured files read by the program. This permits the instructor to directly access the strategies and make modifications to that aspect of the knowledge base.

The inference engine in IP manages the knowledge base and the inferencing process. It governs the questions posed to the instructor and the searching of rules, and tests the goals of the expert system to see if they have been accomplished (Nebendahl, 1988).

The knowledge base of IP is founded on instructional theory that proposes that specific learning conditions require different instructional strategies or strategy modifications in order to optimize learning (Briggs, 1977; Gagné, Briggs, & Wager, 1988; Merrill & Goodman, 1972). IP considers two main categories of learning conditions when determining the best instructional strategy: (1) learning outcome and (2) instructional mode. IP identifies eight levels of learning outcomes and two levels of instructional modes conceptualized as a matrix containing 16 instructional strategies. The eight learning outcomes in IP are as follows:

Response	Paired-associate learning,
Recitation	Memorized verbal information
Explanation	Conceptual understandings of verbal information
Classification	Applying concrete or defined concepts
Prediction	Applying natural laws
Decision	Applying rules of action
Performance	Applying intellectual or physical procedures
Problem Solving	Applying a heuristic process or developing a creative product

The other dimension of the conceptualized matrix is the instructional mode. This refers to whether the lesson is more instructor or learner centered and controlled. The two levels employed are (1) direct instruction and (2) indirect instruction (Borich, 1988). Direct instruction refers to an expository or a more instructor directed approach (Rosenshine, 1986), while indirect instruction refers to a discovery or a more learner directed approach (Harris, Heil & Young, 1983).

The main menu contains the following options:

New Plan	For initializing a new lesson. Here the user can enter the lesson's title and date to be given.
Load Plan	For loading a previously designed and saved plan from any dictionary and drive.
Design Plan	This is the expert system portion of the program. The user answers the multiple choice questions posed by the expert system. The expert system then selects the most appropriate instructional strategy.
Edit Plan	Here the user can add subject matter to a strategy, modify a strategy, create a personalized strategy, etc.
Save Plan	This option is for saving Instructor's Plan files on a directory and rive of choice.
Print Plan	This feature allows the user to print a formatted lesson plan as an ASCII file to a printer or disk.
Quit	Here the user exits the program.

Formative evaluation has shown that instructors find IP very easy to learn to use and they report that it reduces planning time. Instructors who use the program believe its main strength is to improve the quality of their lesson design.References

References

Bahrami, A. (1988). *Designing artificial intelligence based software.* Wilmslow: Sigma Press.

Borich, G. D. (1988). *Effective teaching methods.* Columbus: Merrill Publishing Company.

Briggs, L. J. (1977). The teacher as designer. In L. J. Briggs (Ed.), *Instructional design: Principles and applications.* Englewood Cliffs: Educational Technology Publications.

Gagné, R. M., Briggs, L. J., & Wager W. W. (1988). *Principles of instructional design,* 3rd Ed. New York: Holt, Rinehart, and Winston.

Harris, R. C., Heil, L. H., & Young, J. R. (1983). *Teaching and learning: Psychological and philosophical precepts.* Minneapolis: Burgess Publishing Company.

Levine, R. I., Drang, D. E., & Edelson, B. (1988). *A comprehensive guide to AI and expert systems using turbo pascal.* New York: McGraw-Hill Book Company.

Merrill, M. D., & Goodman, R. I. (1972). *Selecting instructional strategies and media: A place to begin.* National Special Media Institutes.

Nebendahl, D. (1988). *Expert systems: Introduction to the technology and applications.* John Wiley & Sons, Ltd.

Rosenshine, B. V. (1986, April). Synthesis of research on explicit teaching. *Educational Leadership,* 60-69.

Townsend, C., & Feucht, D. (1986). *Designing and programming personal expert systems.* Blue Ridge Summit: Tab Books, Inc.

Wilkins, D. A. (1990). *An introduction to instructor's plan: An expert system for rapid lesson authoring.* A system demonstration at the annual Utah State University Summer Instructional Technology Institute, Logan, Utah, July 26, 1990.

Title:

Cognitive Apprenticeships: An Instructional Design Review

Authors:

Brent Wilson
Peggy Cole

COGNITIVE APPRENTICESHIPS: AN INSTRUCTIONAL DESIGN REVIEW

Brent G. Wilson and Peggy Cole
University of Colorado at Denver

The field of instructional design (ID) emerged more than 30 years ago as psychologists and educators searched for effective means of planning and producing instructional systems (Reiser, 1987; Merrill, Kowallis, & Wilson, 1981). Since that time, instructional designers have became more clearly differentiated from instructional psychologists working within a cognitivist tradition (Resnick, 1981; Glaser, 1982; Glaser & Bassok, 1989). ID theorists tend to concern themselves with prescriptions and models for designing instruction while instructional psychologists conduct empirical research on learning and instructional processes.

Of course, the distinction between designers and psychologists is never clear-cut. Over the years, many psychologists have put considerable energy into the design and implementation of experimental instructional programs. Because the two fields support different literature and theory bases, communication between the two fields is often strained. Thus much design work of cognitive psychologists has gone relatively unnoticed by instructional design theorists.

Collins, Brown, and colleagues (e.g., Collins, Brown, & Newman, 1989) have developed an instructional model derived from an analysis of the way apprentices work under experts in traditional societies and from the way people seem to learn in everyday informal environments (Rogoff & Lave, 1984); they have called their model *cognitive apprenticeships*, and have identified a list of features found in "ideal" learning environments. Instructional strategies, according to the Collins-Brown model, would include modeling, coaching, scaffolding and fading, reflection, and exploration. Additional strategies are offered for representing content, for sequencing, and for maximizing benefits from social interaction.

Of course, many of Collins' recommended strategies resemble strategies found in the instructional-design literature (e.g., Reigeluth, 1983). Clearly both fields could benefit from improved communication concerning research findings and lessons learned from practical tryout. With that goal in mind, the purpose of this paper is to analyze the Collins-Brown cognitive apprenticeship model from an instructional-design point of view. In addition to general strategies and recommended components, several teaching systems employing cognitive apprenticeship ideals are described. The resultant review should prove valuable in two ways: (1) cognitive psychologists should be able to make a better correspondence between their models and current ID theory, hopefully seeing areas needing improvement, and (2) instructional-design theorists also should be able to see correspondences and differences, which may lead to revision or expansion of *their* current models.

The Need for Cognitive Apprenticeships

The cognitive apprenticeship model rests on a somewhat romantic conception of the "ideal" apprenticeship into a complex domain (Brown, Collins, & Duguid, 1989). In contrast to the classroom context, which tends to remove knowledge from its sphere of use, Collins et al. recommend returning instruction to settings where worthwhile problems can be worked with and solved. The need for a problem-solving orientation to education is apparent from the difficulty schools are having in achieving substantial learning outcomes (Resnick, 1989).

Another way to think about the concept of apprenticeship is Gott's (1988) notion of the "lost apprenticeship," a growing problem in industrial and military settings. She noted the effects of the increased complexity and automation of production systems. First, the need is growing for high levels of expertise in supervising and using automated work systems; correspondingly, the need for entry levels of expertise is declining. Workers on the job are more and more expected to be flexible problem solvers; human intervention is often most needed at points of breakdown or malfunction. At these points, the *expert* is called in. Experts, however narrow the domain, do more than apply canned job aids or troubleshooting algorithms; rather, they internalize considerable knowledge which they can use to solve flexibly problems in real time.

Gott's second observation relates to training opportunities. Now, at a time when more problem-solving expertise is needed due to the complexity of systems, fewer on-the-job training opportunities exist for entry-level workers. There is often little or no chance for beginning workers to acclimatize themselves to the job, and workers very quickly are expected to perform like seasoned professionals. True apprenticeship experiences are becoming relatively rare. Gott calls this dilemma—more complex job requirements with less time on the job to learn—*the "lost" apprenticeship*, and argues for the critical need for cognitive apprenticeships and simulation-type training to help workers develop greater problem-solving expertise.

A Brief Review of ID Models

Current instructional-design models are based on Robert Gagne's *conditions-of-learning* paradigm (Gagne, 1985), which in its time was a significant departure from the Skinnerian operant conditioning paradigm dominant among American psychologists. The conditions-of-learning paradigm posits that a graded hierarchy of learning outcomes exists, and for each desired outcome, a set of conditions exists that leads to learning. Instructional design is a matter of being clear about intended learning outcomes, then matching up appropriate instructional strategies. The designer writes behaviorally specific learning objectives, classifies those objectives according to a taxonomy of learning types, then arranges the instructional conditions to fit the current instructional prescriptions. In this way, designers can design instruction to successfully teach a rule, a psychomotor skill, an attitude, or piece of verbal information.

A related idea within the conditions-of-learning paradigm claims that sequencing of instruction should be based on a hierarchical progression from simple to complex learning outcomes. R. Gagne developed a technique of learning hierarchies for analyzing skills: A skill is rationally decomposed into parts and subparts, then instruction is ordered from simple subskills to the complete skill. Elaboration theory uses content structure (concept, procedure, or principle) as the basis for organizing and sequencing instruction (Reigeluth, Merrill, Wilson, & Spiller, 1980). Both depend on task analysis to break down the goals of instruction, then on a method of sequencing proceeding from simple to gradually more complex and complete tasks.

These instructional-design models appear to work with well-defined content domains; their most extensive use has been in military settings. There are a couple of potential problems, however, that need to be addressed. First, sequencing of instruction based on the logical structure of content tends to neglect the existing knowledge base of individual learners. ID models often assume new content can be added incrementally to learners' prior knowledge bases without complication. As a simple example, imagine a concept "kinds" hierarchy of animals, vertebrates, mammals, dogs, collies, etc. Following elaboration theory's method of sequencing concept hierarchies,*dog*, being lower in the hierarchy, is considered a more detailed concept than *mammal*; hence, *dog* should not be taught until the concept of *mammal* has been presented. This prescription, of course, fails to take into account children's existing knowledge structure that includes *dog* at a young age but not *mammal*. A more natural sequencing method would start with *dog* and proceed to the unfamiliar technical concepts of *mammals, vertebrates,* etc.

Another problem with current models is their relative neglect of the job of integrating different content elements in learners' minds. As described above, current models break down content into different types of learning outcomes, and prescribe different learning conditions for each type. This analytic approach builds in a bias toward teaching disconnected, isolated content elements out of their naturally occurring contexts. Because rules or skills are often taught separately from facts or verbal information, their interrelatedness tends to be de-emphasized. This is unfortunate, because rules are best taught in meaning-rich environments, and verbal information is best taught in meaningful contexts that include action and problem-solving performance. In fact, different kinds of outcomes seem to depend on each other for optimal learning environments. Teaching one kind of content in exclusion of other kinds creates problems for both learner and teacher. Elaboration theory addresses the problem by recommending periodic synthesizers and summarizers, but its sequencing strategy centers around one kind of learning outcome (concept, procedure, or principle) for an entire course. Thinking of content structure only in terms of concept hierarchies, procedures, or causal models is overly restrictive in light of cognitive research suggesting the importance of other kinds of knowledge structure (Strike & Posner, *in press*).

Some of the teaching models being offered by cognitive researchers bear strong resemblance to traditional ID models. Larkin and Chabay (1989), for example, offer design guidelines for the teaching of science in the schools (pp. 160–163):

1. Develop a detailed description of the processes the learner needs to acquire.
2. Systematically address all knowledge included in the description of process.
3. Let most instruction occur through active work on tasks.
4. Give feedback on specific tasks as soon as possible after an error is made.
5. Once is not enough. Let students encounter each knowledge unit several times.
6. Limit demands on students' attention.

By any standard, these design guidelines are very close to the prescriptions found in component display theory, elaboration theory, and Gagne's instructional-design theory. The strong correspondence can be seen as good news for instructional-design theories: Current cognitive researchers seem to agree on some fundamentals of design that also form the backbone of ID models.

On the other hand, some teaching models recently developed and tested emphasize design elements that traditional ID models historically have under-emphasized. The cognitive apprenticeship model includes some well-worn design elements—such as modeling and fading—and others that are relatively neglected by ID models—such as situated learning, exploration, and the role of tacit knowledge.

FEATURES OF COGNITIVE APPRENTICESHIPS

In the section below, the design elements of the cognitive apprenticeship model (based primarily on Collins, 1991) are reviewed and related to traditional ID concepts.

1. ***Content: Teach tacit, heuristic knowledge as well as textbook knowledge*** Collins et al. (1989) refer to four kinds of knowledge that differ somewhat from ID taxonomies:

- Domain knowledge is the conceptual, factual, and procedural knowledge typically found in textbooks and other instructional materials. This knowledge is important, but often is insufficient to enable students to approach and solve problems independently.
- Heuristic strategies are "tricks of the trade" or "rules of thumb" that often help narrow solution paths. Experts usually pick up heuristic knowledge indirectly through repeated problem-solving practice; however, heuristic knowledge can be made explicit and taught directly.
- Control strategies are required for students to monitor and regulate their problem-solving activity. Control strategies have monitoring, diagnostic, and remedial components; this kind of knowledge is often termed *metacognition* in the literature (Paris & Winograd, 1990).
- Learning strategies are strategies for learning ; they may be domain, heuristic, or control strategies, aimed at learning. Inquiry teaching to some extent directly models expert learning strategies (Collins & Stevens, 1983).

Comment

ID taxonomies (Gagne, 1985; Merrill, 1983) pertain primarily to domain or textbook knowledge, although the distinction is not explicit. Gagne's "cognitive strategies" fits the last three strategies listed by Collins et al. Merrill's "find a principle" or "find a procedure" seems closest to those three strategies. Both Gagne and Merrill are least specific about instruction of this type, although they both acknowledge its importance.

2. ***Situated learning: Teach knowledge and skills in contexts that reflect the way the knowledge will be useful in real life.*** Brown, Collins, and Duguid (1989) argue for placing all instruction within "authentic" contexts that mirror real-life problem-solving situations. Collins (1991) is less forceful, moving away from real-life requirements and toward problem-solving situations: for teaching math skills, situated learning could encompass settings "ranging from running a bank or shopping in a grocery store to inventing new theorems or finding new proofs. That is, situated learning can incorporate situations

from everyday life to the most theoretical endeavors" (Collins, 1991, p. 122).

Collins differentiates a situated-learning approach from common school approaches:

> *We teach knowledge in an abstract way in schools now, which leads to strategies, such as depending on the fact that everything in a particular chapter uses a single method, or storing information just long enough to retrieve it for a test. Instead of trying to teach abstract knowledge and how to apply it in contexts (as word problems do), we advocate teaching in multiple contexts and then trying to generate across those contexts. In this way, knowledge becomes both specific and general.* (Collins, 1991, p. 123)

Collins cites several benefits for placing instruction within problem-solving contexts:

- Learners learn to apply their knowledge under appropriate conditions.
- Problem-solving situations foster invention and creativity (see also Perkins, 1990).
- Learners come to see the implications of new knowledge. A most common problem inherent in classroom learning is the question of relevance: "How does this relate to my life and goals?" When knowledge is acquired in the context of solving a meaningful problem, the question of relevance is at least partly answered.
- Knowledge is stored in ways that make it accessible when solving problems. People tend to retrieve knowledge more easily when they return to the setting of its acquisition. Knowledge learned while solving problems gets encoded in a way that can be accessed again in problem-solving situations.

Although not cited by Collins and colleagues, two other cognitive teaching models are relevant to the notion of situated learning. Bransford and colleagues (Bransford, Sherwood, Hasselbring, Kinzer, & Williams, 1990) have developed an approach called "anchored instruction." They take a Sherlock Holmes or Indiana Jones videodisc and develop problem-solving activities around incidents in the video. The instruction is thus grounded in a rich *macro-context* that is meaningful and interesting to the learner. They have pointed out many similarities to situated learning and cognitive apprenticeships (The Cognition and Technology Group at Vanderbilt, 1990). The second related model is Spiro's *cognitive flexibility theory* (Spiro, Coulson, Feltovich, & Anderson, 1988; Spiro & Jehng, 1990) which grew out of studies of medical students learning advanced subject matter. Spiro and colleagues found that for students to avoid oversimplifying complex content, they needed to see multiple analogies across multiple contexts. Spiro and colleagues used a series of "mini-cases" to help students a mental model that was sensitive to the many nuances and subtleties of the content. Both Bransford's anchored instruction and Spiro's cognitive flexibility theory stress the importance of placing problem-solving instruction within meaningful contexts.

Comment

Taken in its extreme form that would require "authentic," real-life contexts for all learning, the notion of situated learning is somewhat vague and unrealistic. Instruction always involves the dual goals of generalization and differentiation (Gagne & Driscoll, 1989). In its more modest form, however, the idea of context-based learning has considerable appeal. Gagne & Merrill (1990) have pointed to the need for better integration of learning goals through problem-solving "transactions." The notion of situated learning, however it is viewed, challenges the conditions-of-learning paradigm that prescribes the breaking down of tasks to be taught out of context.

3. ***Modeling and explaining: Show how a process unfolds and tell reasons why it happens that way.*** Collins (1991) cites two kinds of modeling: (1) modeling of processes observed in the world and (2) modeling expert performance, including covert cognitive processes. Computers can be used to aid in the modeling of these processes. Collins stresses the importance of integrating both the demonstration and the explanation during instruction. Learners need access to explanations as they observe details of the modeled performance. Computers are particularly good at modeling covert processes that otherwise would be difficult to observe, both natural and mental. Collins suggests that truly modeling expert performance, including the false starts, dead ends, and backup strategies, can help learners more quickly adopt the tacit forms of knowledge alluded to above under the *Content*

section. Teachers in this way are seen as "intelligent novices" (Bransford, Goin, Hasselbring, Kinzer, Sherwood, & Williams, 1988). By seeing both process modeling and accompanying explanations, students can develop "conditionalized" knowledge, that is, knowledge about when and where knowledge should be used to solve a variety of problems.

Comment

ID models presently incorporate modeling and demonstration techniques. Tying explanations to modeled performances is a useful idea, similar to Chi and Bassok's (1989) studies of worked-out examples. Again, the emphasis is on making tacit strategies more explicit by directly modeling mental heuristics.

4. ***Coaching: Observe students as they try to complete tasks and provide hints and helps when needed.*** Intelligent tutoring systems sometimes embody sophisticated coaching systems that model the the learner's progress and provide hints and support as practice activities increase in difficulty. Burton and Brown (1982) developed a coach to help learners with "How the West Was Won," a game originally implemented on the PLATO system. Anderson, Boyle, and Reiser (1985) developed coaches for geometry and LISP programming.

   Coaching strategies can be implemented at least partially in traditional school settings. Bransford and Vye (1989) identify several characteristics of effective coaches:
   - Coaches need to monitor learners' performance to prevent their getting too far off base, but leaving enough room to allow for a real sense of exploration and problem solving.
   - Coaches help learners reflect on their performance and compare it to others'.
   - Coaches use problem-solving exercises to assess learners' knowledge states. Misconceptions and buggy strategies can be identified in the context of solving problems; this is particularly true of computer-based learning environments (Larkin & Chabay, 1989).
   - Coaches use problem-solving exercises to create the "teaching moment." Posner, Strike, Hewson, and Gertzog (1982) present a 4-stage model for conceptual change: (1) students become dissatisfied with their misconceptions; (2) they come to a basic understanding of an alternative view; (3) the alternative view must appear plausible; and (4) they see the new view's value in a variety of new situations (see also Strike & Posner, *in press*).

Comment

Coaching probably involves the most "instructional work" (*cf.* Bunderson & Inouye, 1987) of any of the cognitive apprenticeship methods. Short of one-on-one tutoring, coaching is likely to be partial and incomplete. Cooperative learning and small-group learning methods can provide some coaching support for individual performance. And computers can help tremendously in monitoring learner performance and providing real-time helps; yet presently coaching is only fully implemented in resource-intensive intelligent tutoring systems. Much work is being done to model the essentials of coaching functions on computer systems; we continue to need resource-efficient methods for achieving the coaching function.

5. ***Articulation: Have students think about their actions and give reasons*** for their decisions and strategies, thus making their tacit knowledge more explicit. Think-aloud protocols are one example of articulation (Hayes & Flower, 1980; Smith, 1988). Collins (1991) cites the benefits of added insight and the ability to compare knowledge across contexts. As learners' tacit knowledge is brought to light, that knowledge can be recruited to solve other problems.

Comment

John Anderson (1990) has shown that procedural knowledge—the kind that people can gain automaticity in—is initially encoded as declarative or conceptual knowledge, but later fades as the skill becomes proceduralized. Methods for articulating tacit knowledge helps to restore a conscious awareness of those lost strategies, enabling more flexible performance. Traditional ID models suggest practicing problem solving to learn problem solving, but are surprisingly lacking in specific methods to teach learners to think consciously about covert strategies.

6. ***Reflection: Have students look back over their efforts to complete a task and analyze their own performance.*** Reflection is like articulation, except it is

pointed backwards to past tasks. Analyzing past performance efforts can also involve elements of strategic goal-setting and intentional learning (Bereiter & Scardamalia, 1989). Collins and Brown (1988) suggest four kinds or levels of reflection:

- *Imitation* occurs when a batting coach demonstrates a proper swing, contrasting it with your swing;
- *Replay* occurs when the coach videotapes your swing and plays it back, critiquing and comparing it to the swing of an expert;
- *Abstracted replay* might occur by tracing an expert's movement of key body parts such as elbows, wrists, hips, and knees, and comparing those movement to your movements;
- *Spatial reification* would take the tracings of body parts and plot them moving through space.

The latter forms of reflection clearly rely on technologies—video or computer— for instantiation. Collins (1991) uses Anderson et al.'s Geometry Tutor as an example of reflective instruction.

Comment

Articulation and reflection are both strategies to help bring meaning to activities that might otherwise be more "rote" and procedural. Reigeluth's (1983) concern with meaningful learning is indicative of the need; however, much of traditional ID practice tends to devalue the reflective aspects of performance in favor of getting the procedure down right.

7. ***Exploration: Encourage students to try out different strategies and hypotheses and observe their effects.*** Collins (1991) claims that through exploration, students learn how to set achievable goals and to manage the pursuit of those goals. They learn to set and try out hypotheses, and to independently seek knowledge. Real-world exploration is always an attractive option; however, constraints sometimes prohibit extensive time in realistic settings. Simulations are one way to allow exploration; hypertext structures are another.

Comment

As Reigeluth (1983) notes, discovery learning techniques are less efficient than direct instruction techniques. The choice must depend, to some extent, on the goals of instruction: for near transfer tasks (*cf.* Salomon & Perkins, 1988), direct instruction may occasionally be warranted; for far transfer tasks, learners must learn not only the content but also now to solve unforeseen problems using the content; in such cases, instructional strategies allowing exploration and strategic behavior become essential.

Having thus represented a traditional ID mode of thinking about exploration, I am still left unsatisfied. Following a cognitive-apprenticeship way of thinking, there is something intrinsically valuable about situated problem-solving activity that makes learning work better than straight procedural practice. This is an area that needs better articulation among ID theorists.

8. ***Sequence: Present instruction in an ordering of simple to complex, increasing diversity, and global before local skills.***
   - *Increasing complexity*. Collins et al. (1989) point to two methods for helping learners deal with increasing complexity. First, instruction should take steps to control the complexity of assigned tasks. They cite Lave's study of tailoring apprenticeships: apprentices first learn to sew drawers, which have straight lines, few pieces of material, and no special features like zippers or pockets. They progress to more complex garments over a period of time. The second method for controlling complexity is through scaffolding; for example, group or teacher support for individual problem solving.
   - *Increasing diversity* refers to the variety in examples and practice contexts.
   - *Global before local skills* refers to helping learners acquire a mental model of the problem space at very early stages of learning. Even though learners are not engaged in full problem solving, through modeling and helping on parts of the task (scaffolding), they can understand the goals of the activity and the way various strategies relate to the problem's solution. Once they have a clear "conceptual map" of the activity, they can proceed to developing greater skill at specific skills.

Comment

The sequencing suggestions above bear a strong resemblance to those of elaboration theory (Reigeluth & Stein, 1983) and component display theory (Merrill, 1983). The notion of global before local skills is implicit in elaboration theory; simple-to-complex sequencing is the foundation of elaboration theory. The notion of increasing diversity is the near-equivalent to the prescription to use "varied example" and practice activities in concept or rule learning. The cognitive apprenticeship extends these notions beyond rule learning to problem-solving contexts.

EXAMPLES OF COGNITIVE APPRENTICESHIP

In the section below, I briefly review three approaches that Collins et al. (1989) identify as embodying cognitive apprenticeship features.

Procedural Facilitations for Writing

Novice writers typically employ a *knowledge-telling* strategy: they think about their topic, then write their thought down; think again, then write the next thought down, and so on until they have exhausted their thoughts about the topic. This strategy, of course, is in conflict with a more constructive, planning approach in which writing pieces are composed in a more coherent, intentional way. To encourage students to adopt more sophisticated writing strategies, Scardamalia and Bereiter (1985) have developed a set of writing prompts called *Procedural Facilitations,* that are designed to reduce working-memory demands and provide a structure for completing writing plans and revisions. Their system includes a set of cue cards for different purposes of writing, structured under five headings: *new idea, improve, elaborate, goals,* and *putting it together*. Each prompt is written on a notecard and drawn by learners working in small groups. The teacher makes use of two techniques, *soloing* and *co-investigation*. Soloing gives learners the opportunity to try out new procedures by themselves, then return to the group for critique and suggestions. Co-investigation is a process of using think-aloud protocols that allow learner and teacher to work together on writing activities. This allows for more direct modeling and immediate direction. Bereiter and Scardamalia (1987) have found up to tenfold gains in learning indicators with nearly every learner improving his/her writing through the intervention.

Reciprocal Teaching

Brown & Palinscar (1989) have developed a cooperative learning system for the teaching of reading, termed *Reciprocal Teaching*. The teacher and learners assemble in groups of 2 to 7 and read a paragraph together silently. A person assumes the "teacher" role and formulates a question on the paragraph. This question is addressed by the group, whose members are playing roles of *producer* and *critic* simultaneously. The "teacher" advances a summary, and makes a prediction or clarification, if any is needed. The role of teacher then rotates, and the group proceeds to the next paragraph in the text. Brown and colleagues have also developed a method of assessment, called *dynamic assessment,* based on successively increasing prompts. The Reciprocal Teaching method uses a combination of modeling, coaching, scaffolding, and fading to achieve impressive results, with learners showing dramatic gains in comprehension, retention, and far transfer over sustained periods.

Schoenfeld's Math Teaching

Schoenfeld (1985) studied methods for teaching math to college students. He developed a set of heuristics that were helpful in solving math problems. His method introduces those heuristics, as well as a set of control strategies and a productive belief system about math, to students. Like the writing and reading systems, Schoenfeld's systems includes explicit modeling of problem-solving strategies, and a series of structured exercises affording learner practice in large and small groups, as well as individually. He employs a tactic he calls "postmortem analysis," retracing the solution of recent problems, abstracting out the generalizable strategies and components. Unlike the writing and reading systems, Schoenfeld carefully selects and sequences practice cases to move learners into higher levels of skill. Another interesting technique is the equivalent to "stump the teacher," with time at the beginning of each class period devoted to learner-generated problems that the teacher is challenged to solve. Learners witnessing occasional false starts and dead ends of the teacher's solution can acquire a more appropriate belief structure about the nature of expert math problem solving. Schoenfeld's positive research findings support a growing body of math research suggesting the importance of acquiring a conceptual or schema-based representation of math problem solving.

Forecasting Apprenticeships

COMET (Cooperative Program for Operational Methodology Education and Training) is an inter-agency program that develops weather training for forecasters affiliated with the National Weather Service, the Air Weather Command, and the Navy. Interactive videodisc training modules developed for these forecasters are also made available to university meteorology departments across the nation. COMET's distance learning program is presently nearing beta testing of its first two modules, *Doppler Interpretation* and *Convection Initiation*. The Collins-Brown cognitive apprenticeship model provides the conceptual foundation for the module series. Each module uses optical laserdisc, CD-ROM, and object-oriented programming to simulate a forecaster workstation and provide forecasting practice under representative conditions. The elements of the cognitive apprenticeship model are discussed below as they relate to the design of the *Convection Initiation* module.

Content. Cognitive apprenticeships teach domain or textbook knowledge, but they also make explicit the strategic knowledge needed to use that knowledge in solving real problems. A key goal for the COMET module was to provide authentic activities that required the forecaster to use situation-specific knowledge to make a forecast. There are no distinct "concept" lessons or "rule" lessons. Instead, a series of forecasting problems are presented with accompanying "conceptual models" and procedural "tutorials." In short, the domain and strategic content is made available to the learner in the immediate context of solving specific forecasting problems.

Methods. Computer-based instruction (CBI) is relatively well-suited to cognitive apprenticeship teaching methods. In addition to the modeling and scaffolding common to CBI, a growing number of toolkit-style programs and simulations allow for degrees of exploration. Reflection—comparing one's own problem-solving processes with others'—is possible though somewhat difficult with CBI. The COMET modules replicate a sophisticated forecasting workstation, where both learner and program may initiate actions toward problem solution, presentation, or feedback. However, because the module lacks a natural language interface, the interaction is somewhat constrained. A continuing challenge for CBI design is to develop methods of making instruction approach a continuous dialogue rather than a forced-choice, prefabricated path.

Sequence. As noted above, the sequencing rules of a cognitive apprenticeship may depend on the type of content being taught—e.g., hypertext exploration for schema building (Wilson & Jonassen, 1989) versus simple-to-complex case selection for skill building (*cf*. Reigeluth & Stein, 1983). The *Convection Initiation* module utilizes a progression of forecasting cases in a simple-to-complex order, similar to Schoenfeld's approach to math teaching; at the same time, a variety of hypertext structures—data, rules, definitions, etc.—are available to the learner on demand using hypertext access structures. This is to be expected: often a mix of well-structured problems and hypertext-like exploration is probably optimal for problem-solving instruction.

Sociology. The social aspects of the cognitive apprenticeship model seem to pose the greatest challenge for CBI design. From what we know about learning, social variables are powerful mediators in learning (e.g., Eckert, 1990; Salomon, Globerson, & Guterman, 1989). CBI is often administered on an individual basis, with a single learner engaged at a computer workstation. Although group work is a desirable option, sometimes it is not possible. The COMET weather forecasting modules are a case in point. Modules are to be completed on the job, but because forecasting offices are sometimes staffed with only 2 - 3 persons on duty at a time, learners must complete instruction individually during break times while others continue on the job. The design challenge is: how can we build in needed social reinforcement and convey a sense of community in what is essentially an individualized learning experience? To address the cognitive apprenticeship's social design, the *Convection Initiation* module incorporated several strategies, including:

1. The module begins with a forecasting office scenario. The learner is teamed with "Ron," a fellow forecaster, and immediately given a forecasting problem to solve. This same office scenario is used as a wrap-up, where the learner completes a 2-hour forecasting shift containing four forecasts.
2. Following the beginning office scenario, the learner is briefly introduced to a

panel of three forecasting experts. These experts serve as guides through the module.

3. Following each forecasting practice, one of the three experts provides feedback on the case in the form of an "expert answer." This is a 30 - 90 second explanation of the case. Meteorological data on-screen are highlighted with arrows; this is combined with either a CD audio overlay or audio-video sequence of the expert using a chromaboard, much like a TV weatherperson explaining a map. The module's beta testing will be completed over the spring and summer; we will report different effects of audio-only vs. chromaboard explanations. In either case, the personalized feedback from the actual experts should allow some social modeling to occur. The meteorological field is fairly small; moreover, there are no annual conventions typically attended by members of the field. Getting to know the forecasting experts through the IVD modules should help strengthen the sense of community within the field.
4. As stated above, the module concludes with an office-based simulation where the forecaster, teamed with "Ron," solves a series of realistic, time-based forecasting problems.

Through these efforts to make the module more personalized, the social elements needed to support good instruction are artificially created. We believe these compensating strategies will help improve the reception of the modules in the field, and improve the attitudes of learners completing the modules on an individual basis.

SUMMARY AND CONCLUSION

This paper has been concerned with the relationship between two related disciplines: cognitive psychology and instructional design. ID, the more applied discipline, faces the challenges of constantly re-inventing itself as new research in basic psychology sheds lights on learning processes. It is no cause for concern that ID models need continual self-review; indeed, there would be cause for concern if the situation were otherwise. To provide some sense of context, however, I would like to call attention to some basic principles of cognitive apprenticeship articulated by the philosopher Erasmus more than 500 years ago. In a letter to a student friend, Erasmus offers some advice:

> *Your first endeavor should be to choose the most learned teacher you can find, for it is impossible that one who is himself no scholar should make a scholar of anyone else. As soon as you find him, make every effort to see that he acquires the feelings of a father towards you, and you in turn those of a son towards him .... Secondly, you should give him attention and be regular in your work for him, for the talents of students are sometimes ruined by violent effort, whereas regularity in work has lasting effect just because of its temperance and produces by daily practice a greater result than you would suppose .... A constant element of enjoyment must be mingled with our studies so that we think of learning as a game rather than a form of drudgery, for no activity can be continued for long if it does not to some extent afford pleasure to the participant.*
>
> *Listen to your teacher's explanations not only attentively but eagerly. Do not be satisfied simply to follow his discourse with an alert mind; try now and then to anticipate the direction of his thought .... Write down his most important utterances, for writing is the most faithful custodian of words. On the other hand, avoid trusting it too much .... The contests of minds in what we may call their wrestling ring are especially effective for exhibiting, stimulating, and enlarging the sinews of the human understanding. Do not be ashamed to ask questions if you are in doubt; or to be put right whenever you are wrong.*
>
> Erasmus (1974, pp. 114–115)

At a time when terms like *metacognition*, *scaffolding* , and *cognitive apprenticeship* are being invented to describe the learning process, it is humbling to be reminded of the insights of former paradigms. Instructional-design theorists may chafe at the continuing need to revise their theories in light of advances in psychological theory, but it is good for both fields for the dialogue to continue. Indeed, the interaction between basic psychology and applied instructional design can be expected to continue for some time to come.

REFERENCES

Anderson, J. R. (1990). ***Cognitive psychology and its implications*** (3rd ed.). New York: Freeman.

Bereiter, C., & Scardamalia, M. (1987). ***The psychology of written composition***. Hillsdale, NJ: Erlbaum.

Bereiter, C., & Scardamalia, M. (1989). Intentional learning as a goal of instruction. In L. B. Resnick (Ed.), ***Knowing, learning, and instruction: Essays in honor of Robert Glaser***. Hillsdale, NJ: Erlbaum.

Bransford, J. D., Goin, L. I., Hasselbring, T. S., Kinzer, C. K., Sherwood, R. D., & Williams, S. (1988). Learning with technology: Theoretical and empirical perspectives. ***Peabody Journal of Education, 64*** (1), 5–26.

Bransford, J. D., Sherwood, R. D., Hasselbring, T. S., Kinzer, C.K., & Williams, S. M. (1990). Anchored instruction: Why we need it and how technology can help. In D. Nix & R. Spiro (Eds.), ***Cognition, education, and multimedia: Exploring ideas in high technology***. Hillsdale, NJ: Erlbaum.

Bransford, J. D., & Vye, N. J. (1989). A perspective on cognitive research and its implications for instruction. In L. B. Resnick & L. E. Klopfer (Eds.), ***Toward the thinking curriculum: Current cognitive research***. 1989 ASCD Yearbook.

Bunderson, C. V., & Inouye, D. K. (1987). The evaluation of computer-aided educational delivery systems. In R. M. Gagne (Ed.), ***Instructional technology: Foundations***. Hillsdale, NJ: Erlbaum.

Collins, A. (1991). Cognitive apprenticeship and instructional technology. In L. Idol & B. F. Jones (Eds.), ***Educational values and cognitive instruction: Implications for reform***. Hillsdale, NJ: Erlbaum.

Collins, A., & Brown, J. S. (1988). The computer as a tool for learning through reflection. In H. Mandl & A. Lesgold (Eds.), ***Learning issues for intelligent tutoring systems***. New York: Springer.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and arithmetic. In L. B. Resnick (Ed.), ***Knowing, learning, and instruction: Essays in honor of Robert Glaser***. Hillsdale, NJ: Erlbaum.

Collins, A., & Stevens, G. (1983). Inquiry teaching. In C. M. Reigeluth (Ed.), ***Instructional-design theories and models: An overview of their current status***. Hillsdale, NJ: Erlbaum.

Eckert, P. (1990). Adolescent social categories—Information and science learning. In J. Gardner, J. G. Greeno, F. Reif, A. H. Schoenfeld, & A. A. diSessa, & E. Stage (Eds.), ***Toward a scientific practice of science education***. Hillsdale, NJ: Erlbaum.

Erasmus, D. (1974). ***The correspondence of Erasmus***. Volume 1: Letters 1 to 141, 1484 to 1500. Translated by R. A. B. Mynors and D. F. S. Thomson. Toronto: University of Toronto Press.

Gagne, R. M. (1985). ***The conditions of learning*** (4th ed.). New York: Holt, Rinehart, & Winston.

Gagne, R. M., & Driscoll, M. P. (1989). ***Essentials of learning for instruction*** (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Gagne, R. M., & Merrill, M. D. (1990). Integrative goals for instructional design. ***Educational Technology Research and Development, 38*** (1), 23–30.

Glaser, R. (1982). Instructional psychology: Past, present, and future. ***American Psychologist, 37*** (2), 292–305.

Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology, 40*, 631–666.

Gott, S. P. (1988, April). ***The lost apprenticeship: A challenge for cognitive science***. Paper presented at the meeting of the American Educational Research Association, New Orleans.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Sternberg (Eds.), ***Cognitive processes in writing***. Hillsdale, NJ: Lawrence Erlbaum Associates.

Merrill, M. D., Kowallis, T., & Wilson, B. G. (1981). Instructional design in transition. In F. Farley & N. Gordon (Eds.), ***Psychology and education: The state of the union.*** Chicago: McCutcheon.

Palinscar, A. M., & Brown. A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. **Cognition and Instruction**, 1, 117 - 175.

Paris, S. G., & Winograd, P. (1990). How metacognition can promote academic learning and instruction. In B. F. Jones & L. Idol (Eds.), ***Dimensions of thinking and cognitive instruction***. Hillsdale, NJ: Erlbaum.

Perkins, D. N. (1990). The nature and nurture of creativity. In In B. F. Jones & L. Idol (Eds.), ***Dimensions of thinking and cognitive instruction***. Hillsdale, NJ: Erlbaum.

Posner, G. K., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982) Accommodation of scientific conception: Toward a theory of conceptual change. ***Science Education**, 66,* 211–227.

Reigeluth, C. M. (Ed.). (1983). ***Instructional-design theories and models: An overview of their current status.*** Hillsdale, NJ: Erlbaum.

Reigeluth, C. M. Meaningfulness and instruction: Relating what is being learned to what a student knows. ***Instructional Science,*** 12(3) 197–218.

Reigeluth, C. M., & Stein, R. (1983). Elaboration theory. In C. M. Reigeluth (Ed.), ***Instructional-design theories and models: An overview of their current status.*** Hillsdale, NJ: Erlbaum.

Reigeluth, C. M., Merrill, M. D., Wilson, B. G., & Spiller, R. T. (1980). The elaboration theory of instruction: A model for structuring instruction. ***Instructional Science**, 9,* 195 - 219.

Reiser, R. A. (1987). Instructional technology: A history. In R. M. Gagne (Ed.)., ***Instructional technology: Foundations.*** Hillsdale, NJ: Erlbaum.

Resnick, L. B. (1981). Instructional Psychology. **Annual Review of Psychology.**

Resnick, L. B. (1989, February). Learning in school and out. ***Educational Researcher,*** 5 - 8.

Rogoff, B. (1990). ***Apprenticeship in thinking: Cognitive development in social context.*** New York: Oxford University Press.

Rogoff, B., & Lave, J. (Eds.). (1984). ***Everyday cognition: Its development in social context*** (pp. 95 - 116). Cambridge, MA: Harvard University Press.

Salomon, G., Globerson, T., & Guterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a reading partner. ***Journal of Educational Psychology,*** 81(4), 620 - 627.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. ***Educational Psychologist**, 24*(2), 113-142.

Scardamalia, M., & Bereiter, C. (1985). Fostering the development of self-regulation in children's knowledge processing. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), ***Thinking and learning skills: Research and open questions*** (pp. 563 - 577). Hillsdale, NJ: Erlbaum.

Schoenfeld (1985). ***Mathematical problem solving.*** New York: Academic Press.

Smith, P. (198). Think-aloud protocol paper.

Spiro, R. J., Coulson, R. L., Feltovich, P. J., & Anderson, D. K. (1988, October). ***Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains*** (Technical Report No. 441). Champaign, IL: University of Illinois at Urbana-Champaign, Center for the Study of Reading.

Spiro, R. J., & Jehng, J-C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Eds.), ***Cognition, education, and multimedia: Exploring ideas in high technology*** (pp. 163-205). Hillsdale, NJ: Lawrence Erlbaum Associates.

Strike, K. A., & Posner, G. J. (in press). A revisionist theory of conceptual change. To appear in R. Duschl & R. Hamilton (Eds.), ***Philosophy of science, cognitive science, and educational theory and practice.*** Albany, NY: SUNY Press.

The cognition and technology group at Vanderbilt. (1990, August-September). Anchored instruction and its relation to situated cognition. ***Educational Researcher,*** 2–10.

Wertsch, J. V. (1989). ***Vygotsky and the social formation of mind.*** Cambridge, MA: Harvard University Press.

Wilson, B. G., & Jonassen, D. H. (1989). Hypertext and instructional design: Some preliminary guidelines. ***Performance Improvement Quarterly,*** *2*(3), 34 - 49.

Wood, D. J. (1980). Teaching the young child: Some relationships between social interaction, language, and thought. In D. R. Olson (Ed.), ***The social foundations of language and thought.*** New York: Norton.

Title:

A Qualitative Study of Teachers' Planning of Instruction for Adult Learners in a Telecommunications-based Distance Education Environment

Author:

Linda L. Wolcott

A QUALITATIVE STUDY OF TEACHERS' PLANNING OF INSTRUCTION FOR ADULT LEARNERS IN A TELECOMMUNICATIONS-BASED DISTANCE EDUCATION ENVIRONMENT

by

Linda L. Wolcott, Ed.D.

According to Schön (1983), the context for much of today's professional practice constitutes "not problems to be solved, but problematic situations characterized by uncertainty, disorder, and indeterminacy" (p. 15-16) in which "problems do not present themselves to practitioners as givens" (p. 40). Teaching at a distance presents educators with such a unique situation of practice.

With the growth of telecommunications technology, distance education has entered a new generation. Today, degree programs, training, and continuing professional education are often delivered to adult learners by telecommunications technologies such as telephone, television, computer, and satellite. These media allow learners who are physically separated from their instructor, and often from one another, to transcend geographical distance. By providing for real-time participation and immediate two-way interaction among participants, the application of these telecommunications media creates a unique form of instruction and a unique situation of practice.

Distance instruction differs contextually from traditional classroom instruction. Mediating instruction via telecommunications technology to overcome physical distance alters both the interpersonal and instructional communications processes. Consequently, the "joint activities of teaching and learning at a distance raise a completely new set of problems" (Sparkes, 1983, p. 183). One such "problem" often confronting distance teachers is that of designing instruction for adult learners.

Faced with new delivery technologies, altered teaching environments, and changing student demographics, distance teachers must "accomodate [sic] a number of variables which are not normally encountered" (Carter, 1982, p. 6) in designing instruction for more conventional settings. Planning instruction for a distance context becomes problematic. Those who design and present instruction for telecommunicated distance delivery to adults face new challenges in designing instruction to assure that desired outcomes are achieved.

This study addressed the lack of an understanding drawn from practice of how teachers of college and adult students design instruction for live, telecommunicated delivery. Given that teachers do not apply prescriptive planning models (Morine-Dershimer & Vallance, 1976; Neale, Pace & Case, 1983; Peterson, Marx & Clark, 1978; Sardo, 1982; Taylor, 1970; Yinger, 1977; Zahorik, 1975), what constitutes planning for distance instruction for adult learners? What constructs and rules guide distance teachers in designing instruction in a complex and uncertain practice setting?

The purpose of this study was to describe the preactive instructional planning of university faculty teaching adult learners at a distance. Four questions shaped the course of the study:

1. What is the nature of teachers' planning?
2. What instructional considerations most concern distance teachers as they plan?
3. What are the major factors that influence teachers in planning distance instruction in their particular setting?
4. What rules of action governing planning can be inferred from teachers' self-reports of planning and from observations of their implemented plans?

Theoretical Foundation

Jackson (1968) distinguished two components of teaching: interactive and preactive teaching. Interactive teaching refers to those phases when the teacher is working directly with the students. Preactive teaching occurs prior to and after interactive teaching when teachers are engaged in activities such as preparing lessons, selecting instructional materials, or evaluating students' performance. During both of these component phases, teachers plan. Instructional planning is an important aspect of the practice of teaching (Romberg, 1980). In planning at the preactive level teachers set their direction, their course of action. Teachers make decisions about what to teach and how to teach it.

Research in the relatively new area of teacher planning (see Clark & Peterson, 1986; and Shavelson & Stern, 1981) provides insights into how teachers in elementary schools plan. However, the literature tells us little about how teachers of college and adult students plan. Studies are only beginning to examine the instructional planning process in which teachers of college students engage (Andresen et al., 1985; Dinham, 1989; Powell & Shanker, 1982; Stark et al., 1988). Planning for non-traditional learners and environments is unchartered territory.

Unfortunately, the technical/professional knowledge of practice may not provide sufficient guidance for those designing distance instruction (Sparkes, 1983). Further, the literature that pertains to designing instruction to be delivered live at a distance is deficient. While offering practical suggestions and recommendations, the literature fails to provide data-based and conceptually-linked design prescriptions to inform practice. Given these inadequacies, practitioners must look to their own reflective practice.

Professionals, in general, hold implicit theories of practice through which they make sense of complex, divergent situations (Argyris & Schön, 1974). Research in the emerging area of teachers' thinking (see Clark & Peterson, 1986) documents that teachers, as professionals, possess tacit understandings of their roles, their students, and the processes of instruction that guide their practice (Conners, 1978; Elbaz, 1981; Fox, 1983; Marland, 1977; Munby, 1983). However, researchers have not as yet identified rules of action that derive from examining the practice of instruction as it is planned and delivered by telecommunications in a distance context. It remains unknown how faculty, when faced with the uncertain and divergent situation of distance teaching, construct their planning/design practice as a problem that they can solve.

Methods

This study was framed as a qualitative study that described how distance teachers planned instruction for adult learners in a telecommunications-based distance education setting. Planning was examined within the context of instruction that takes place live, is mediated by telecommunications technology, and serves an adult audience.

Participants

Eleven full-time university faculty members who were teaching credit courses at a distance participated in the study. All participants had prior experience in or preparation for teaching at a distance, and had reputations as good teachers. The faculty were teaching courses that were part of the standard curriculum and that they had taught previously in a face-to-face setting. The majority of the students enrolled in the course were adult learners. A profile of participating faculty is presented in Table 1.

Setting

The distance education program linked students at sites located in the states of Utah and Colorado with live instruction originating from a classroom at a state university in Utah. From eight to eleven remote receive sites were connected by telecommunications technology to one of two distance classrooms at the origination site.

Instruction was delivered by audio-graphic media that involved the transmission of two-way audio and one-way still image video signals via telephone lines. Audio communication was facilitated by the use of push-to-talk microphones. Communication of visual images was accomplished in several ways. Both originating classrooms had the capacity to transmit still video (slow-scan/freeze-frame) images of the instructor, of writing on the board, or of a prepared visual such as an overhead transparency. Classroom A used an electronic blackboard to transmit real-time writing; Classroom B was equipped with an Optel writing tablet that also enabled the transmission and annotation of previously stored graphics. As many as eleven sites widely dispersed over the two states depended on daily UPS deliveries for the submission and return of assignments and tests.

Data Collection

Data were collected over the course of the winter and spring quarters of 1990. Three data collection methods were applied: interview, observation, and document analysis. The primary method of data collection was semi-structured interviewing. Each of the faculty members was interviewed twice during the quarter in which s/he was teaching. The interviews emphasized the process of planning instruction. Questions were structured around three general themes: the activities of planning; instructional concerns considered during planning; and influences that affected planning. Each interview was audio-taped and subsequently transcribed into an interview log for analysis.

Second, the researcher directly observed at least two class sessions conducted at a distance by each of the faculty interviewed. Since the phenomenon of planning could not be observed directly, observation sought to derive a sense of context for which instruction was planned. All observations were conducted in the originating classroom. The researcher made detailed field notes, focusing on both instructional presentation and dynamics.

Table 1

Research Participants

Participant	Department	Rank	Distance Teaching. Exp.	Type of Course	Class-room	Number of Sites	Number of Students
1	Psychology	Assoc. Prof.	Veteran	Lower Div.	A	11	50
2	Anthropology	Professor	Veteran	Upper Div.	B	7	18
3	History	Professor	Veteran	Upper Div.	B	10	37
4	Business	Professor	Veteran	Upper Div.	A	10	66
5	English	Professor	New	Upper Div.	B	8	46
6	Education	Asst. Prof.	Veteran	Graduate	B	9	45
7	English	Asst. Prof.	New	Upper Div.	B	7	18
8	Economics	Assoc. Prof.	Veteran	Graduate	B	7	18
9	English	Assoc. Prof.	Veteran	Upper Div.	B	9	39
10	Philosophy	Assoc. Prof.	Veteran	Upper Div.	B	5	35
11	Business	Professor	Veteran	Lower Div.	A	9	34
					Average	8	37

The third means of data collection was the analysis of documents pertaining to instruction as planned by the participating teachers. The researcher collected two types of documents from the participants: (1) pre-existing planning documents such as lesson plans, presentation outlines, or course syllabi, and (2) a journal generated specifically for the study.

Data Analysis

In analyzing the data, the researcher applied the constant comparative method of data analysis. Constant comparison involved coding, categorizing, and theorizing about the data. Data from each interview, observation and document were coded to yield descriptive and interpretive categories that were compared with data collected subsequently. As data were compared, categories were merged or revised, and properties of those categories were developed. The final phase in the constant comparative method involved theorizing, that is, finding relationships among the categories.

To facilitate the handling of large amounts of text data, QUALPRO, software designed expressly for the management of qualitative data, was used. The program worked in conjunction with word processing software, (in this case WordPerfect 5.0), to store, code, organize and manipulate text files.

Findings

The results of data analysis describe the process in which faculty engaged when planning, the major factors that influenced planning, and the teacher-held principles--implicit rules of action--that faculty applied in planning. Table 2 summarizes the categories of data analysis and their respective properties as described below.

The Nature of Planning

Three features characterized the preactive planning process: (1) faculty engaged in course or term planning, (2) planning was driven by content, and (3) distance planning focused on the development of an extended syllabus.

Term planning. When designing distance instruction, the participants approached the task by planning at the course or term level. Planning took the form of a time-consuming, front-end activity rather than an ongoing one. Looking at the big picture, the faculty saw their course of action laid out over the duration of the quarter. They concentrated their planning efforts on making the majority of their

Table 2

Summary of Findings

The Nature of the Planning Process
- Term planning
- Driven by content
- Focused on syllabus development

Factors that Influence Planning
- Constrained by time
- Restricted by the medium
  - Lack of visual communication
  - Physical separation
  - Technical obstacles
  - On-site students
  - Lag time

Implicit Rules of Action
- Cover the material
- Maintain control
- Go with the flow
- Provide an equitable learning experience

instructional decisions up front where the course was essentially packaged in an extensive syllabus prior to the beginning of the term. There was little day-to-day lesson planning.

In planning the course, the participants were concerned, first, with defining the content and, second, with matching the content with the time available in which to teach it. These two tasks were embedded in the larger, central planning activity, the development of the course syllabus.

How to present content went largely unplanned. By their own admission, faculty were secondarily concerned with methodology, as evidenced by the lack of written planning documents other than the syllabus and by the absence of procedural plans within the syllabi themselves.

Driven by content. To these distance teachers planning instruction meant planning content. Participants centered on the selection and sequencing of the subject matter. In both the decisions they made and in the planning documents they produced, the faculty focused on content--that "set of information" that one selected "to cover". While a few participants mentioned considering factors such as the characteristics of their "audience", or the influence of logistical and technical constraints on instruction, it was a preoccupation with what to teach that drove the instructional design of the course.

In planning the course, two dependent tasks dominated: determining "what to cover", and matching content with the time available in which to cover it. Determining what to cover consisted of identifying the concepts and setting goals and/or objectives. Defining content was a process of giving shape to an amorphous body of knowledge.

Faculty asked themselves the "what" questions that gave shape to the content: "What's the course for?"; "What do I really want students to accomplish at the end of this class?"; "What are the abstractions that they have to understand?"; "What's important and what's kind of frosting on the cake?"

The second planning task, matching content, involved establishing the order of coverage and fitting content to the given time frame. Matching time and content meant "equat[ing] the time you have available to the material you want to cover." Once they defined the relevant content, typically in a topic outline or lecture notes, the faculty fit or matched it to the blocks of time available.

Focused on syllabus development. Decisions about content were packaged in an extensive syllabus prior to the first meeting of the term. The syllabus embodied the planning decisions and captured them in a tangible form. As one participant described it, "developing the syllabus is where it's at. . . Any course you've taught for a while has its shape in your head, and the syllabus is nothing more than giving it form, a concrete form." Developing a syllabus was the focal point of distance planning.

To call this central planning document a syllabus is misleading; for in most cases, the syllabus was far more extensive than the traditional one- or two-page handout. It was commonly referred to as the "expanded", "enhanced", or "extended" syllabus, and in its extension, the syllabus was unique.

The extended syllabus had as its base the traditional class syllabus which typically contained items such as a course description; listings of goals, readings and assignments; a topic outline; and grading policies. In addition to these standard items, the syllabi also contained handouts or hard copies of visual materials, study questions, reprinted articles, or extensive essays. Syllabi varied widely in length, contents and format. A summary of syllabus contents is presented in Table 3. It was the marriage of the standard document with expanded lecture notes and supplemental materials that gave birth to the new expanded form that was predominantly a content document.

The syllabus served several purposes. First, it was designed to support the semi-independent learning of the students. Since most distance classes met less frequently, and some for shorter periods of time than those which met on campus face-to-face, an enhanced syllabus compensated for the loss of contact hours. Particularly for those distance students who did not have ready access to the instructor or to a peer learning group, the syllabus was a valuable aid that directed their thinking and learning.

Providing an enhanced syllabus also economized instructional time: "Because the syllabus is so much beefier in terms of reading material and exercises that I normally expect in an on-campus course, than I can actually reduce the amount of lecture time." The syllabus was, at the same time, as much a study guide for students as it was a master plan for the instructor. A number of faculty used the syllabus as a presentation outline relying on its organizational structure to guide their lectures. In class, the syllabus served as "a point of departure for discussion" from which the teacher and the students could "flesh out the details by talking together.

Table 3
Syllabus Contents

	Features																				
Participant	course description	course goals	topic outline / schedule	reading assignments	assignment description	grading policy	textbook	address / phone no.	office hours	lesson objectives	content narrative	articles / readings	study questions	optional reading list	handouts / hard copies	glossary	vocabulary lists	exercises	"how to study" tips	length of syllabus pp.	Special Features
1		x	x	x		x	x	x	x				x		x		x			89	
2		x	x	x	x		x	x		x	x			x				x		46	session outline
3	x	x		x	x	x	x	x			x		x				x	x	x	77	
4		x	x	x		x	x	x	x			x			x					28	
5	x	x	x	x	x	x	x	x	x		x	x		x		x			x	61	
6	x	x	x	x	x	x	x		x	x	x	x			x					136	session outline
7	x	x	x	x		x	x	x	x	x	x	x	x		x				x	267	sample quiz
8	x		x	x	x	x	x	x						x						3	
9			x	x	x	x	x	x	x	x	x		x		x				x	30	
10	x	x	x	x		x	x	x		x	x	x		x	x	x			x	156	illustrated biographical sketches space for notes outlines of arguments
11			x	x		x	x								x		x	x		33	solutions to problems

Producing the syllabus also represented professional development. Syllabus development was time-consuming and called for reflecting on customary approaches to teaching, as a veteran distance teacher observed: "I don't know of anybody who sits down and thinks through a class the way this system makes you think through it to crank out a sixty or seventy page syllabus."

Developing the syllabus had a reciprocal effect on instruction. An interesting spin-off of this central planning activity was that a number of faculty used the extended syllabus with subsequent sections of the same course taught entirely in a face-to-face setting. Additionally, faculty reported that the extra planning efforts of producing the syllabus paid off in improved teaching.

But developing and working from an extended syllabus had an important drawback. Faculty reported feeling "tied to the syllabus". On the one hand, an extended syllabus gave the course its structure and laid out its direction. Yet once written, the syllabus became, in effect, a contract leaving little room for spontaneous deviation. Faculty felt locked in to following the syllabus, and consequently, less free to be flexible or spontaneous.

Factors That Influenced Planning

Designing distance instruction was subject to the influence of factors in the teaching and learning environment. Two attributes of the distance setting, time and the medium of delivery, defined the context in which the instruction would take place. These two contextual factors influenced teachers in planning their instructional course of action, and affected the outcome of planning.

Constrained by time. Time was a significant factor in designing distance instruction. Typically, there were fewer contact hours for the distance taught courses than for the same course offered on campus. The humanities courses, for example, met once a week for two hours when the comparable on-campus course met for one hour four times a week. Teaching in this distance program meant working within a compacted or condensed time frame in which faculty expected to cover the same material but in less time.

Accommodating a foreshortened time frame involved both compromise and dilemma. Time constraints had the effect of diminishing spontaneity. Given the fewer contact hours coupled with the brevity of the quarter system, most faculty expressed a sense of urgency in accomplishing their goals. In classrooms, the atmosphere was often tense as participants felt "up against the clock", particularly when technical or logistical difficulties further robbed them of instructional time.

Participants often spoke of time as a luxury. Contrasting distance teaching to face-to-face teaching, participants noted that in the distance context, time was a precious commodity to be judiciously consumed. Deviating from planned content or permitting the instructional exchange to develop unencumbered constituted a luxury they could not afford.

Restricted by the medium of delivery. In their planning, the participants were also influenced by the medium of delivery. The medium had the effect of restricting what the teacher could plan to do, both because of how the system was configured, and because of the logistical challenges such a configuration posed. It was common in journal entries for the writers to vent their irritation with logistical and technical obstacles and interruptions that frustrated their plans. Likewise, conversations with participants invariably touched upon problems and annoyances associated with the mechanics of the program that impinged upon what an instructor planned or would have planned to do.

Five dimensions of the medium and its delivery emerged as challenges to designing distance instruction.

(1) the lack of visual communication - The inability to see the students during instruction was perceived as a major "electronic barrier" to instructional communication that impeded feedback, rapport, and interaction. Without the return video, the participants felt that the amount of feedback they were able to provide and, more important, receive from students decreased significantly. Trying to teach without seeing students was tantamount to "teaching in a vacuum", a prospect that caused anxiety in most faculty who had come to depend on feedback to direct the flow of instruction.

Coincidentally, without the opportunity to make eye contact and to gather feedback, teachers had the tendency to refer to students as "faceless people", "invisible faces", or collectively and impersonally as "out there." They found it difficult "to build a rapport with the class if you can't reach out and touch them" or "to demonstrate personal interest in them when they're a voice out of a black box." The faculty also felt that lacking visual communication, students were less likely to engage in instructional dialogue with them and other students they could not see. And, indeed, only rarely was student-to-student interaction observed.

(2) physical separation - Much the same was said with regard to the second challenge, the physical separation of learners. The amount of interaction among students and between the students and the teacher was reduced along with the spontaneity of such dialogue. Participants compared the dynamics between conventional and distance settings: "When you have an on-campus group, they work off of each other. . . The [distance] students [are] without that kind of interaction in the class."

Again, rapport was hindered as teachers found it difficult to get to know the students, and for the students to get to know students other than those with whom they shared a site. Teaching and learning at a distance represented not only physical distance among those involved, but also "psychic distance".

(3) technical obstacles - The configuration of the system and technical difficulties often associated with its implementation affected the dynamics of instruction, and the use of instructional materials. Problems with equipment or transmission could lessen the teaching value of the instructional material, or at worst, render a particular instructional material useless. Technical failures were particularly troublesome, for they broke the "flow" or natural development of planned instruction and spontaneous discussion. Interruptions represented "a pause in the discussion. . . a mental pause for me".

Frustrations with the implementation of instructional materials discouraged their use. While teaching at a distance begged instructors to "think visually", the realities prompted faculty to ask whether investing time and energy in planning for the use of visuals was worth the effort.

(4) on-site students - The presence (or absence) of students at the origination site posed a fourth challenge to design. The majority of participants preferred to have students in the originating classroom. In addition to providing a personal dimension that many found essential to their teaching style, "real people to whom I could relate" contributed a critical element to the dynamics of instruction. Teachers relied on the face-to-face students to provide them with feedback about their teaching performance and to "gauge" the students' comprehension. Without the feedback provided by "the presence of warm bodies in front of you. . ." , "it becomes really, really easy just to lecture". Though they preferred to have some students present, the participants acknowledged an

inherent risk: the tendency to teach to the students they could see and with whom they could directly interact.

(5) lag time - A fifth influence on design was logistical, and partly bureaucratic. A major challenge existed in overcoming "lag time" between the receipt and redistribution of assignments and tests. Configured to accommodate one-way video and two-way audio, alternative means were required for the exchange of printed materials. Options such as facsimile were costly and its frequent use discouraged. The mail, while more economical, took more time and was subject to the vagaries of distribution on campus and at the sites. The logistics of this system posed a bottleneck that reduced timely feedback to many students.

The factors of time and the medium of delivery--sometimes alone, often in combination, exerted their influence on the design of distance instruction. Each set limits that challenged teachers' planning.

Rules of Action: Implicit Guides to Planning

The concern for content and the factors of time and the medium, were not the sole determinants of instruction. Faculty also brought to planning their beliefs, attitudes, and concerns about the nature and conduct of instruction. The participants developed a set of guiding principles or implicit rules of action that blended important elements from their own teaching philosophies, styles and experiences with the factors at work in distance planning. The following tacit rules underlaid teachers' planning.

Cover the material. Once unburdened of what to cover, faculty were guided by the directive to "cover the material". The defined content became the instructor's "agenda". Because of "time pressures", adhering to that agenda drove the implementation of instruction.

Following the precept to cover material, the faculty designed instruction that was tightly structured and predominantly lecture. Being structured meant having a clear sense of what the objectives were and where those objectives should lead. Writing an extended syllabus gave the course its structure and provided both faculty and students with the course and session agendas. Preoccupation with covering content lead participants to utilize an expository mode of presentation. "Lecture with some discussion" was the primary instructional method.

The choice to cover the material involved trade-offs. Most faculty members reported that "engaging" students was important to their teaching style, but in reality interaction took a back seat to content as the participants frequently sacrificed discussion and dialogue to covering the material. But as an English professor told his class that's "the price you pay for distance learning".

Maintain control. Facing barriers to communication and intrusions on their classroom autonomy, faculty were motivated by the need to be in control of the instruction. In planning, one way to maximize one's control was through structure. The teacher determined and controlled the agenda that was tightly centered around content and written in a detailed syllabus. Strategies employed to maintain control resulted in instruction that was teacher-centered and predominantly verbal.

While content was under the faculty's control, the delivery and the physical environment were not. To give themselves more of a sense of control, the participants played it safe and minimized their dependence on the system. They used familiar instructional techniques which, for most, were a combination of lecture and questioning, and they stayed close to the agenda. Playing it safe also meant using familiar instructional media such as the chalkboard, overhead transparencies or an occasional video tape. The visual capabilities of the medium were under-utilized; prepared graphics such as slides and computer generated graphics were rarely used. To explore the capabilities of the new medium of delivery was to risk the lose of control to technical difficulties or logistical delays. By minimizing one's dependence on the system, a teacher could exercise some control within a precarious medium of delivery.

Go with the flow. Participants were also influenced in designing instruction by the concept of "the flow of instruction". Flow referred to the natural or "organic" development of instruction that brought together planned activities and spontaneous discussion. Achieving flow depended on the instructor being flexible enough to allow spontaneous deviation from plans, structured enough to know where the instruction should lead, and in control enough to be able to bring digressions back on course. One participant defined it as "a balance that is difficult to keep between allowing students to ask a lot of questions, to engage in a lot of interaction, and, quote 'covering the material'". In going with the flow, the faculty applied flexibility within structure. Having content structured afforded them the opportunity to be responsive and flow with the instruction, but also to bring the direction of instruction back to their set agenda.

Premised on the assumption that instructional interaction flowed naturally, generated by "the quality and interest of the material" and "provocative questions", faculty modeled their plan on that of on-campus instruction where students more readily took a participatory role and dialogue was more spontaneous. In a traditional campus environment, interaction developed rather "organically" because it was not imperative to build in opportunities for interaction and feedback. At a distance, however, and with obstacles to free-wheeling communication, the distance students were frequently less active contributors to or initiators of discussion. Though the faculty desired and encouraged student participation, they did not deliberately plan for interaction. Unplanned, interaction was dependent on skillful on-line questioning and the capricious nature of student participation. As a result, the teacher shouldered the responsibility for nurturing flow.

Assuming that flow would develop naturally often proved counterproductive. Without sufficient interaction and feedback, and in the absence of planned contingencies to stimulate flow, the faculty found themselves less able to "orchestrate the flow of instruction" and more inclined to lecture.

Provide an equitable learning experience. Participants were also guided by the concern to provide distance students with, as much as possible, the same learning experience as the on-campus students were afforded. Sensitive to impressions among some of their colleagues that distance learning was a "second best system" of instruction, faculty expressed the need to provide for an equal experience. The principle of providing an equitable learning experience translated into doing things the same way.

Concerns for equity reinforced the use of familiar methods and media, and discouraged thinking about distance teaching on its own terms. It was common in comparing contexts for participants to conceptualize distance teaching as an anomaly, referring to classes taught at a distance and their "normal" or "regular" class. The challenge of planning distance instruction was to find ways to pursue the same ends by the same means.

Discussion

Although teaching at a distance prompted subtle differences in style, participants concluded that their approach to instruction remained essentially the same in a distance setting as it was in a more conventional environment. But in process, however, it was the creation of a unique planning document that distinguished planning for distance instruction from designing its on-campus counterpart.

Developing the extended syllabus was synonymous with planning the distance course, for in its development, the syllabus embodied the decisions faculty made about the shape of instruction. The extended syllabus was, at once, the focal point and manifestation of distance planning. The creation of a detailed syllabus supports Stark's et al. (1988) speculation that syllabus development, in general, "may represent the process of course development in which the teacher has engaged" (p. 17).

The emphasis accorded the development of a detailed syllabus in the present context, strongly suggests its continued evolution and applicability to other settings. While the syllabus took on added significance for these teachers because of a compressed instructional time frame, the development of some type of teaching/learning document is likely to be a major preoccupation of most distance teachers' planning.

While faculty were preoccupied with determining the content to be covered, and matching that content to the time available, their approach to the design of instruction glossed over a number of general design considerations. Factors such as the circumstances under which the instruction would take place, how instruction would be delivered, and the intended learning outcomes received little attention in the faculty's construction of their course of action.

Especially absent from their expressed planning concerns were considerations of the students. That is not to say that faculty failed to take students into account when planning, but the minimal amount of talk about students' needs, interests, and learning attributes suggested that students were a much lower priority in a planning scheme that focused on content. Even more noticeably, the faculty did not talk about adult students, who constituted the greater part of the learner population. The majority of faculty did not acknowledge adults as learners distinct from the more traditional university student. The slighting of adult students resulted not from a conscious, deliberate choice to ignore them, but rather from a lack of awareness of the characteristics and needs of this population, and of the implications for teaching and learning.

Separated by distance and relying principally on audio communication, teachers were challenged to think visually and to maximize interactivity. However, the instruction, as designed by the participants, was not characterized by either of these attributes. Rather, distance instruction looked much like traditional face-to-face college instruction: it was teacher-centered lecture and assumed a natural emergence of student participation and involvement. Little, if any, content was carried or supplemented graphically; interactive teaching and learning techniques were not the norm. The faculty missed an opportunity to optimally integrate the capabilities of the delivery medium into the instructional design.

When considering the instructional concerns and influences that are represented in the teachers' implicit guides to planning distance instruction, one conclusion stands out above the others. In planning distance instruction, faculty retrofitted rather than redesigned their instruction. That is, the participants did not so much design or plan instruction in consideration of the content, the student characteristics, and the context of instruction as they did adapt or transform instruction that had previously been planned. Redesigning would have called for accommodating the contextual factors and the characteristics of learners by building an instructional plan from the ground up with respect to content and method.

In retrofitting, by contrast, participants made adjustments in style or adaptations in strategies to fit a previously existing course and its methods to a new context. The planning problem became one of making the minimal changes required to fit customary instruction to a different environment.

Retrofitting had the effect of mirroring face-to-face instruction. The faculty began with existing content and an existing course design. They conceded to the constraints of time and the restrictions imposed by the medium and, adjusted instruction as designed for face-to-face presentation to fit the distance context. Changes came about in response to restrictions on customary procedures, rather than in consideration of a changed environment. The strategy of retrofitting amounted to transferring instruction from a face-to-face setting to a distance medium rather than transforming its design.

Of the courses taught at a distance during the present study, all were pre-existing courses, that is, the faculty member had taught the course previously in a traditional campus setting. None were both new to the instructor and new to the distance mode of delivery. Given the time pressures and faced with the novelty of distance teaching, faculty found it both safe and expedient not to change. But what if the courses had been planned from scratch, specifically to be taught at a distance? Would faculty have designed them differently? The question remains open to speculation.

This discussion does not mean to imply that retrofitting or teachers' customary practices were wrong, and that designing from scratch would be the only way to approach distance instruction. Nor is it to say that instruction was not effective as designed and implemented. The final outcome of instruction may well have been the same. It is, however, the conviction of this researcher that designing distance instruction requires giving full consideration to the elements that have the potential to shape the design. Retrofitting as a design approach can be detrimental. A mind-set for adapting face-to-face instruction hinders faculty in closely reflecting on the course design that best reconciles the content with the students' characteristics and with the assets and liabilities of the medium of delivery. As Knapper (1988) contends, designing distance instruction in terms of its on-campus counterpart is a disservice.

Conclusion

This study looked at the preactive planning of a small group of faculty in one particular distance teaching and learning context. Features of the instructional environment such as the reduced contact hours, the amount of training and instructional support provided to faculty, and the audio-graphic delivery system, restrict the transfer of findings to other distance contexts. However, the findings expand the literature on teachers' thinking by providing additional insight into the little explored area of how college/university faculty plan. But more than that, this description of planning in higher education ventures into two areas not touched upon by previous research. The present study inaugurates research on planning in non-traditional settings with non-traditional learners.

REFERENCES

Andresen, L., Barrett, E., Powell, J., & Wieneke, C. (1985). Planning and monitoring courses: University teachers reflect on their teaching. Instructional Science, 13, 305-328.

Argyris, C., & Schön, D. (1974). Theory in practice: Increasing professional effectiveness. San Francisco: Jossey-Bass.

Carter, J.F. (1982, March). Instructional design for distance training. Paper presented at the annual meeting of the American Educational Research Association, New York. (ERIC Document Reproduction Service No. ED 217 214)

Clark, C.M., & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock (Ed.), Handbook of research on teaching (3rd ed.). (pp. 255-296). New York: Macmillan.

Conners, R.D. (1978). An analysis of teacher thought processes, beliefs, and principles during instruction. Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Dinham, S.M. (1989). College teachers' thinking and planning: A qualitative study in the design studio. (ERIC Document Reproduction Service No. ED 313 373)

Elbaz, F. (1981). Teacher's practical knowledge: Report of a case study. Curriculum inquiry, 11, 43-71.

Fox, D. (1983). Personal theories of teaching. Studies in Higher Education, 8, 151-163.

Jackson, P.W. (1968). Life in the classroom. New York: Holt, Rinehart and Winston.

Knapper, C. (1988). Lifelong learning and distance education. American Journal of Distance Education, 2(1), 63-72.

Marland, P.W. (1977). A study of teachers' interactive thoughts. Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Morine-Dershimer, G., & Vallance, E. (1976). Teacher planning. (Beginning Teacher Evaluation Study Special Report C). San Francisco: Far West Laboratory for Educational Research and Development.

Munby, H. (1983, April). A qualitative study of teachers' beliefs and principles. Paper presented at the annual meeting of the American Educational Research Association, Montreal. (ERIC Document Reproduction Service No. ED 228 215)

Neale, D.C., Pace, A. J., & Case, A.B. (1983, April). The influence of training, experience, and organizational environment on the teacher's use of the systematic planning model. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec.

Peterson, P.L., Marx, R.W., & Clark, C.M. (1978). Teacher planning, teacher behavior, and student achievement. American Educational Research Journal, 15, 417-432.

Powell, J.P., & Shanker, V.S. (1982). The course planning and monitoring activities of a university teacher. Higher Education, 11, 289-301.

Romberg, T.A. (1980). Salient features of the BTES framework of teacher behaviors. In C. Denham & A. Lieberman, Time to learn (pp. 73-93). Sacramento: State Commission for Teacher Preparation.

Sardo, D. (1982, October). Teacher planning styles in the middle school. Paper presented at the meeting of the Eastern Educational Research Association, Ellenville, NY.

Schön, D. (1983). The reflective practitioner: How professionals think in action. New York: Basic Books.

Shavelson, R.J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. Review of Educational Research, 51, 455-498.

Sparkes, J.J. (1983). The problem of creating a discipline of distance education. Distance Education, 4, 179-186.

Stark, J.S., Lowther, M.A., Ryan, M., Bomotti, S.S., Genthon, M., Martens, G., & Haven, C.L. (1988). Reflections on course planning: Faculty and students consider influences and goals. Ann Arbor: National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan.

Taylor, P.H. (1970). How teachers plan their courses: Studies in curriculum planning. Slough, Berkshire, England: National Foundation for Educational Research in England and Wales.

Yinger, R.J. (1977). A study of teacher planning: Description and theory development using ethnographic and information processing methods. (Doctoral dissertation, Michigan State University). Dissertation Abstracts International, 38, 207A-208A.

Zahorik, J.A. (1975). Teachers' planning models. Educational Leadership 33, 134-39.

Title:

An Anthropological View of Educational Communications and Technology: Beliefs and Behaviors in Research and Theory

Author:

Andrew R. J. Yeaman

An Anthropological View Of Educational Communications And Technology: Beliefs And Behaviors In Research And Theory[1]

Andrew R. J. Yeaman
Address: Yeaman & Associates
601 West Eleventh Avenue #1103
Denver, CO 80204
Telephone: (303) 534-5749

Abstract: This paper examines the possibility of a Kuhnian paradigm in research and theory in educational communications and technology. First, it analyzes Kuhn's model and the attraction of this model as a viewfinder. It turns to anthropology for explanation of beliefs and behaviors. Then it takes a critical look at use of the Kuhnian paradigm metaphor in the field of educational communications and technology. The implications of an alternative research agenda are examined. This paper concludes with a discussion of the schism between researchers and practitioners in the ways they know the world.

The paradigmatic status of educational communications and technology is reviewed in this paper from a cultural anthropology perspective. Kuhn's metaphorical book on paradigms, *The Structure of Scientific Revolutions* (1970), has itself become a metaphor since its first publication in 1962. Kuhn used citation analysis, a historical method, for building his paradigm model of how empirical thought develops in physics and astronomy. Empiricists from other disciplines borrow his model but they do not apply his methodology. Using rational tools they would not normally allow in their work due to lack of empirical rigor, they describe their own fields as paradigmatic. The Kuhnian metaphorical structure has diffused from describing scientific research in physics and astronomy. It is applied to almost any perceived state or desired change, from improvements in computer operating systems to research and theory in educational communications and technology.

Kuhn's Model

Kuhn's *The Structure of Scientific Revolutions* (1970) opposed Popper's principle of falsifiability (Popper, 1934/1968). Described in *The Logic of Scientific Discovery,* falsification had been accepted for three decades. Popper asserted the superiority of empirical observation over scientific theories accepted on the basis of agreement between authorities. In the Popperian tradition, competition between research strategies is thought advantageous to science. The credibility of steadystate knowledge rests not on dogma but on refinement and replacement by more powerful theories and closer approximations of truth. Kuhn questioned what had become the textbook explanation of continual, logical progress and created his model to explain the results of bibliographic research on the history of science.

---

[1] This paper first appeared in the *Canadian Journal of Educational Communication*, Vol. 19, No. 3, pages 237-246.

Kuhn's model contains three central metaphors: paradigm, anomaly and revolution (1970). The paradigm is an accepted set of rules for knowing about and conducting normal science. The anomaly is the exception that stimulates new explanations that cannot be ignored. The revolution is the emergence of a new paradigm.

Kuhn identified three normal foci of factual scientific investigation (1970, p. 25-30):

Determination of significant fact–Paradigmatic facts are developed for solving paradigmatic problems as a result of applying research strategies. Measurement occurs with increasingly refined apparatus and methods. Some researchers receive more recognition for developing research tools than for what they find.

Matching facts with theory–This comes from addressing research issues. Increasing the match between theory and nature comes from arguments and/or factfinding demonstrations rooted in the real world. Kuhn stated: "The existence of the paradigm sets the problem to be solved; often the paradigm theory is implicated directly in the design of apparatus able to solve the problem." (1970, p. 27).

Articulating theory–A paradigm is articulated by looking for universal constraints, quantitative laws and experiments, in a more qualitative sense, that elucidate a phenomenon and relieve ambiguous interpretations.

Kuhn's Model In Action

This model with its triple foci is not Popperian falsification (Kuhn, 1970, p. 77-80). The historical study of scientific development shows that new basic theories appear in the short revolutionary period of extraordinary science and new paradigms are incommensurable with previous paradigms. In the long periods of normal science, the reigning paradigm restricts researchers to puzzle solving science in which assumptions are accepted and not questioned.

Kuhn (1970) further challenged modern rationality by doubting the neutrality of investigators and suggested the importance of considering sociological and individual psychological aspects. He situated the longterm search for objective truths within the reality of the everyday world: "Lifelong resistance, particularly from those whose productive careers have committed them to an older tradition of normal science, is not a violation of scientific standards but an index to the nature of scientific research itself." (p. 151).

Kuhn's relativistic approach led to debates with other philosophers and historians of science including Popper, who defended the neutrality of scientists, Lakatos, who described rules for sequential theories in scientific research programs, and Feyerabend, who welcomed subjectivity and accepted mysticism (Lakatos & Musgrave, 1970). The choice of empiricism as a way of knowing drew doubts because it is empirically unprovable. Dependency on gathering data through observation became understood as a bias. Science in action claims objectivity but fails any test of neutrality. Debate with Lakatos chased empiricism into a corner as a sociopolitical business with sociopolitical aims when Kuhn demonstrated the importance of consensus in the scientific community for determining facts.

Kuhn's bibliographic analyses increased in particularity over time because each revolution requires a unique explanation. His original metaphors did not fit all situations. He discovered differing paradigms for chemistry and physics that described helium either as an atom or as a molecule (1970, p. 50-51). Kuhn drew parallels from science to art, from theories to painting styles (1977, p. 340-351). With the refinement of specificity, Kuhn's model was reduced in scale from paradigms in conflict to theories in conflict. This admitted lack of generality, partly in response to close examinations of the issues, resulted in severe attack. Stegmüller's book *The Structure and Dynamics of Theories* finds "Kuhnianism" not only relativistic but irrational (1976, p. 241-246).

Kuhn had constructed a usable metaphor. The model became popular as scholars in many areas borrowed his structure regardless of its relativistic base (belief in social perspective as reality). It has been applied uncritically by realists (believers in objective reality) and instrumentalists (believers in measurement as reality). Kuhn's model has become dominant. Casti wrote, "With Kuhn we have come to the end of the line as far as contemporary views on the way science operates both to form and to validate its view of the world." (1989. p. 45).

Paradigm has lost its revolutionary fire. It may mean no more than *Weltanschauung* as a metaphysical, epistemological and methodological perspective of the times. As the author of a classic work, Kuhn has endured the fate of classic authors because his model has been more often cited than read (Adams & Searle, 1986, p. 381).

Kuhn's Model And Social Science

In examining the relationship of Kuhn to the field of educational communications and technology, it is first necessary to review Kuhn's influence on the social sciences, the source of education theories and research methods. The acknowledgement of a paradigm is socially desirable in any discipline as a sign of intellectual adulthood. Just as early psychological researchers had "physics envy" (Gould, 1981, p. 262-263, some social sciences have been the subject of debates over whether they are really scientific or not (Kuhn, 1970, p. 160). Psychologists claim that understanding paradigm shifts in their field is central to understanding cognitive psychology (Lachman, Lachman & Butterfield, 1979). There was a Kuhnian diffusion of ideas in linguistics:

> In accordance with Thomas Kuhn's (1970) description of paradigm changes in the sciences, the Chomsky point of view took over, not by convincing the previous generation it had been in error, but by winning the allegiance of the most gifted students of the succeeding generation. (Gardner, 1986, p. 209).

Similarly, parapsychologists are attracted by the orthodoxy of Kuhnian metaphors (Barnes, 1983, p. 90-93; Radner & Radner, 1982, p. 62–67). However, Kuhn had ventured "...it remains an open question what parts of social science have yet acquired such paradigms at all. History suggests that the road to a firm research consensus is extraordinarily arduous." (1970, p. 15).

Despite Kuhn avoiding extrapolation of his ideas to the social sciences, the central argument that more is happening in science than an academic competition of ideas caused Barnes to write on *T.S. Kuhn and the Social Sciences* (1983). Barnes describes the dangers of Whig history, of viewing the past as a reflection of the present, and of writing textbooks that present only facts that support current understandings. He writes that Kuhn's ideas have become progressively more conformist, conservative and supportive of the scientific establishment. Barnes also uses the phrase "intellectual laziness" (p. 120) to show that he does not endorse the dogmatic acceptance of Kuhn's model as an after the fact explanation in sociology, economics or psychology.

Kuhn (1970) had cautioned:

> The members of all scientific communities, including the schools of the "pre-paradigm" period, share the sorts of elements which I have collectively labeled 'a paradigm.' What changes with the transition to maturity is not the presence of a paradigm but rather its nature. Only after the change is normal puzzle-solving research possible. (p. 179).

Barnes' (1983) reflections on Kuhn (1970) illustrate the seductiveness of Kuhn's model. No area wants to be regarded as preparadigmatic when having a paradigm appears to be a measure of social standing. The widespread use of Kuhn's model causes a subtle dislocation. There is a self contradiction in employing it to assert professional status. Mitchell, an English professor, identified the essence of this interdisciplinary borrowing:

> We can always tell which of two crafts outranks the other by looking at its lexicon. Jargon only runs downhill. You will notice that although educators have borrowed "input" from the computer people, the computer people have felt no need to borrow "behavioral objectives" or "preassessment" from the educators. (Mitchell, 1979, p. 106).

Cultural Evolution

Kuhn, the historian of science, describes the scientific way of knowing but does not provide a sufficient explanation of the sociopolitical forces driving that way of knowing. Cultural anthropologists, however, specialize in that type of problem. They describe what people do and what they say they do and construct explanations for beliefs and how they change. Harris, in particular, has proposed a theory of cultural evolution that accounts for how beliefs and behaviors are formed in response to environmental pressures (1968, 1974, 1977, 1980, 1989). This theory is known as cultural materialism. Harris' theory explains why beliefs and behaviors are shaped by fundamental issues such as food supply and population growth (1985). He has also extrapolated this theory to hyperindustrial life (1987).

Harris gives the basic principle of cultural materialism with these words: "The etic behavioral modes of production and reproduction probabilistically determine the etic behavioral domestic and political economy, which in turn probabilistically determine the behavioral and mental emic superstructures." (1980, p. 55-56).

Etic operations are independently verifiable: "The test of the adequacy of etic accounts is simply their ability to generate scientifically productive theories about the causes of sociocultural differences and similarities." (Harris, 1980, p. 32).

In contrast, emic operations give native informants absolute status in determining the reality, meaningfulness or appropriateness of analyses. These etic and emic distinctions are not mere synonyms for behavioral and mental. They combine into "four objective operationally definable domains in the sociocultural field of inquiry" (Harris, 1980, p. 38).

An example comes from Harris' fieldwork with farmers in Kerala, on the western side of the Indian peninsula (1980, p. 32-40). From the etic view the feeding of male calves is restricted so the gender ratios of the cattle are adjusted through starvation. This suits the local ecological and economic conditions for farming. From the emic view, no farmer would violate the Hindu prohibition against slaughter. The Kerala farmers say that male cattle are weaker, sicker and inherently eat less than female cattle.

The paradoxical relationship between etic and emic views is testable by a cultural comparison. Hindu farmers in parts of India with different local ecological and economic conditions, such as the inland states of the north, value the traction capabilities of cattle. In Uttar Pradesh, the seat of Hindu religion and culture, the mortality rate of cows is significantly higher than that of oxen. Besides slowly starving the female calves, unwanted animals are sold to Moslem traders. Again, the death of the animals because of gender appears intentional (Harris, 1980; Harris, 1985).

This knowledge can be applied to the culture of educational communications and technology research and theory. From the cultural materialist viewpoint the behavioral and mental emics of a culture are determined by the etic forces. The emic projections or reconstructions become beliefs. Harris lists the mental and emic components as conscious and unconscious cognitive goals, categories, rules, plans, values, philosophies and beliefs about behavior (1980, p. 54). Scholarly beliefs are also emic representations and there is a tendency to favor low grade emic stories over high grade etic information (Price, 1980). Travers describes similar myth building behavior about research in education (1987). The next section of this paper looks at the emic superstructure of educational communications and technology.

The Emic Functions Of A Cognitive Paradigm

The claim to a cognitive paradigm in educational communications and technology (Clark & Salomon, 1986; Clark & Sugrue, 1988; Heinich, 1970; Winn, 1989) can be read as a social text. Harris' theory suggests that belief in the cognitive paradigm performs a social function. Like the boost in agricultural production from a Kwakiutl chief redistributing wealth at a potlatch (Harris, 1974), it encourages cooperation in the joint productive effort. Having a paradigm is an indication of being established. The transition from preparadigm state to postparadigm state is widely perceived as the passage of puberty for any discipline and from the viewpoint of cultural evolution, claiming a paradigm has adaptive value. It helps people obtain and maintain employment. When prospective colleagues say they are believers, they increase their chances of survival in the job market. Publishing manuscripts that look outside of cognition in examining what the field does and why it is done, causes schisms and these reduce the centralized power. Conflict is discouraged because it decreases material growth.

To claim a cognitive paradigm impresses other big men[2] such as granting agencies. It reassures school district superintendents and corporate directors of instructional systems that learning is knowable and predictable. Everything appears under control and the scholars who support a cognitive paradigm promise to bend their research efforts to everyone's benefit.

Belief in the cognitive paradigm in educational communications and technology may exist without paradigmatic consensus. Instead of a paradigm, Kuhn's structure may offer another explanation which fits the field better. In Kuhn's model, preparadigmatic research is characterized by the atheoretical factfinding characteristic of prescientific times (1970, p. 15-16). Explanations of phenomena are inadequate. Data are too dense for decoding. Details are missed which are later considered important. In preparadigmatic research, technology is the name given to solving practical problems systematically. Technology parents science.

Understanding The Transmission Of Culture

From the cultural materialist point of view, individuals and groups in the field can still use their reason to choose what they want to believe and what they want to do. Although lacking evidence for this occurring earlier, Harris suggests that deliberate choices may be the only hope for the planet in the face of the ecological emergency (1987, p. 181-183). Researchers in the field could search to make the questions they ask more meaningful, to make their results more useful and for new methodologies and theories. From this viewpoint, sociocultural research represents a strength of educational communications and technology's position as an applied field. Scholars investigating this dimension would recognize more is at stake in educational communications and technology than achievement. Their work would be closer to practice. Driven by sociocultural research issues, their investigations would draw from the theories and methods of the social sciences and the humanities. Some would write in what Husen (1988) identifies as humanism, educations' other way of knowing, as opposed to neopositivism/logical empiricism. These scholars would be concerned about critically understanding cultural reproduction.

Nichols, for example, believes the field might turn from the mechanical study of achievement and select the direction of Habermasian morality (Habermas, 1984; Habermas, 1987; Wells, 1986):

> Education should function, via communicative action, to help us competently reach understanding with one another (the cultural function), fulfill appropriate societal norms (the social function), and develop our personalities (the socialization function), and in the process, learners become involved with objective, practical and emancipatory forms of knowledge. (Nichols, 1989, p. 351).

---

[2] This is an anthropological term denoting leaders who work extremely hard at motivating their followers to be productive. See, for example, Harris (1989, p. 359).

The Schism Between Researchers And Practitioners

The beliefs and behaviors of educational communications and technology are rooted in the practice of using and producing educational media but researchers focus on one set of activities and practitioners focus on another set of activities. Researchers investigate to write scholarly research reports but educational cinematographers investigate to create films and videos. Classroom teachers, school media coordinators and instructional systems developers select images to convey the world to learners. They know the field is effective because educational media employ rhetoric and the way of saying something changes what is said. Their productions are lyric, dramatic and epic. Even computer screens are alive with metaphors.

Adams' *Philosophy of the Literary Symbolic* (1983) confronts the tension between the literary and the scientific ways of knowing: "The war between poetry and philosophy has extended from before Plato's time into our own." (p. 389). These words also apply to educational communications and technology. More than a paradigm, debates between researchers and between researchers and practitioners ensure conscious decisions. These are necessary to defeat the material pressures on beliefs. More than a paradigm, the field needs this conflict between the philosophy of research and the poetry of practice.

As an applied field, the Kuhnian paradigm of revolutionary process does not fit educational communications and technology and neither does the falsificationist myth of orderly progress. The description of the preparadigmatic state fits best. There are material pressures for *claiming* a cognitive paradigm. That claim is an emic fact, whether it is empirically true or not.

References

Adams, H. (1983). Philosophy of the literary symbolic. Tallahassee, FL: Florida State University Press.

Adams, H. & Searle, L. (1986). Critical theory since 1965. Tallahassee, FL: Florida State University Press.

Barnes, B. (1983). T.S. Kuhn and social science. New York: Columbia University Press.

Casti, J. L. (1989). Paradigms lost: Images of man in the mirror of science. New York: William Morrow.

Clark, R. E. & Salomon, G. (1986). Media in teaching. In M. C. Wittrock (Ed.), Handbook of research on teaching: Third edition: A project of the American Educational Research Association (pp. 464-478). New York: Macmillan.

Clark, R. E. & Sugrue, B. M. (1988). Research on instructional media. In D. P. Ely, B. Broadbent & R. K. Wood (Eds.), Educational media and technology yearbook: 1988 (pp. 19-36). Englewood, CO: Libraries Unlimited.

Gardner, H. (1986). The mind's new science: A history of the cognitive revolution. New York: Basic Books.

Gould, S. J. (1981). The mismeasure of man. New York: W. W. Norton.

Habermas, J. (1984). The theory of communicative action: Volume 1: Reason and the rationalization of society. Boston, MA: Beacon Press.

Habermas, J. (1987). The theory of communicative action: Volume 2: Lifeworld and system: A critique of functionalist reason. Boston, MA: Beacon Press.

Harris, M. (1968). The rise of anthropological theory: A history of theories of culture. New York: Thomas Y. Crowell.

Harris, M. (1974). Cows, pigs, wars and witches: The riddles of culture. New York: Vintage.

Harris, M. (1977). Cannibals and kings: The origins of cultures. New York: Vintage.

Harris, M. (1980). Cultural materialism: The struggle for a science of culture. New York: Vintage.

Harris, M. (1985). The sacred cow and the abominable pig: Riddles of food and culture. New York: Simon & Schuster.

Harris, M. (1987). Why nothing works: The anthropology of daily life. New York: Simon & Schuster.

Harris, M. (1989). Our kind: Who we are: Where we came from: Where we are going. New York: Harper & Row.

Heinich, R. (1970). Technology and the management of instruction. Washington, DC: Association for Educational Communications and Technology.

Husen, T. (1988). Research paradigms in education. In J. P. Keenes (Ed.), Educational research, methodology and measurement: An international handbook (pp. 17-20). Oxford, U.K.: Pergamon.

Kuhn, T. S. (1970). The structure of scientific revolutions: Second edition, enlarged. Chicago IL: University of Chicago Press. (Original work published 1962)

Kuhn, T. S. (1977). The essential tension: Selected studies in scientific tradition and change. Chicago, IL: University of Chicago Press.

Lachman, R., Lachman, J. L. & Butterfield, E. C. (1979). Cognitive psychology and information processing: An introduction. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lakatos, I. & Musgrave, A. (1970). Criticism and the growth of knowledge. Cambridge, U.K.: Cambridge University Press.

Mitchell, R. (1979). Less than words can say. Boston, MA: Little, Brown & Company.

Nichols, R. G. (1989). Reconciling educational technology with the lifeworld: A study of Habermas' theory of communicative action. Proceedings of selected research paper presentations at the 1989 Annual Convention of the Association for Educational Communications and Technology, 341-355. (ERIC Document Reproduction Service No. ED 308805).

Popper, K. R. (1968). The logic of scientific discovery. London: Hutchinson. (Original work published 1934)

Price, B. J. (1980). The truth is not in accounts but in account books: On the epistemological status of history. In E. B. Ross (Ed.). Beyond the myths of culture: Essays in cultural materialism (pp. 155-180). New York: Academic Press.

Radner, D. & Radner, M. (1982). Science and unreason. Belmont, CA: Wadsworth.

Stegmüller, W. (1976). The structure and dynamics of theories. New York: Springer-Verlag.

Travers, R. M. W. (1987). Myths, scripts and educational research. Journal of Research and Development in Education, 20 (3), 14-26.

Wells, S. (1986). Jürgen Habermas, communicative competence, and the teaching of technical discourse. In C. Nelson (Ed.). Theory in the classroom (pp. 245-269). Urbana, IL: University of Illinois Press.

Winn, W. (1989). Toward a rationale and theoretical basis for educational technology. ETR&D, 37, 35-46.

Title:

Sociocultural Aspects of Computers in Education

Author:

Andrew R. J. Yeaman

Sociocultural Aspects of Computers in Education

Andrew R. J. Yeaman
Address: Yeaman & Associates
601 West Eleventh Avenue #1103
Denver, CO 80204
Telephone: (303) 534-5749

Educational practice is largely unexamined by the scholars who study the technological delivery of educational messages. The contextualizing sociocultural aspects are especially neglected. Educational communications and technology research concentrates on the efficiency of learning stimuli, instructional techniques and matching these with student characteristics. Manipulation is assumed to be a desirable activity and prediction is assumed to be a desirable goal. Ellsworth comments:

> Such research privileges questions of how individuals process information from educational media over and above questions about the ideological work of meaning construction, content selection, setting of objectives, and media use. The concept of information processing as developed within cognitive psychology and applied in educational technology cannot explain meaning, intent and significance. (1989, p. 49)

However, educational practice is quite suitable for analysis from critical theory viewpoints, as in the studies collected by Apple and Weis (1983). Cherryholmes' book *Power and Criticism* endorses the application of critical theory, "Criticism provides important opportunities to break with dominant readings and interpretations." (1988, p. 158). Cherryholmes warns this valuable quality of criticism may also cause discomfort in people because it demystifies myths. A false rhetoric that makes cultural things appear natural is often confronted (Barthes, 1973, p. 153-155).

Who is Doing What to Your Computer?

The local supermarket has a program for shoppers to save their receipts for putting computers into schools. Fund raising programs have existed in the U.S. for a long time but there is a qualitative difference between past and present efforts. Putting audiovisual equipment in schools in earlier decades differs in social meaning from putting computers in schools today.

The standard models of educational communications represent the linear transmission of information and ignore the question of social context. This raises poststructural questions of power and meaning: Who is doing what to whom? Why are they doing it? What are the material pressures? What are the beliefs? Which people have a voice?

Through these poststructural questions this paper supports the goal of freeing students and teachers from hidden domination (Gibson, 1986; Giroux, 1983, 1988; Giroux & Simon, 1989). Exposing the ideological aspects of computers in school and society to critical analysis gives this papers' audience new beliefs to consider. Students of this audience will subsequently receive instruction that is more self-liberating. The crucial problem is not overcoming teacher resistance to computers or student resistance to computers but the need to teach resistance. In this sense, resistance is redefined (Giroux, 1983):

> Theories of resistance provide a study of the way in which class and culture combine to offer outlines for a cultural politics. Central to such a politics is a semiotic reading of the style, rituals. language, and systems of meaning that constitute the cultural field of the oppressed. (p. 101)

Analysis of Computers in Education as a Cultural Field

The following observations about the sociocultural aspects of computers in education were collected as samples for analysis from the poststructural critical position.

- The computer industry exists to make money for itself. The values of the computer industry support the boardroom mandate to increase stockholders' share value. Hardware and software manufacturers sell products to educators to make a profit.
- A decrease in computer prices created widespread availability. In 1991 an external 40 megabyte hard disk usually costs under $400.00. In 1982 a popular Tandy 33 megabyte hard disk system sold for $14,980.00.
- The computer industry perpetuates itself by targeting the middle class as customers for computers at school and at home (Apple, 1988, p. 166-170).
- The human values of computers as cultural objects can be further corrupted by anthropomorphism. This is the fallacy of relating to computers as though they are people. It accompanies the oppression of relating to people as though they are machines, something described in detail by Zuboff's *In the Age of the Smart Machine* (1988).
- The difficulties of using and learning to use computers are seldom formally exposed. The difficulties are never acknowledged in instructional materials and the time needed for learning is concealed. Nevertheless, the authors of computer manuals show their own reactions to using and learning to use computers in their national organization's magazine. Even the people who foist computer jargon on the manual reading public find it hard to swallow, making it into puns and malapropisms: "Give me some modem cookies" (So, You Think You Know "Technology"?, 1988).
- People fight back against computerization. A mild response by office workers to the difficulties of using and learning to use computers is photocopying and circulating satirical illustrations from unknown sources. Beyond samizdat publication of office humor, computers represent real and alleged threats to health and safety (Bradley, 1989; Scalet, 1987; Stellman & Henifin, 1989).
- Computers in education are being pushed as a solution but teachers have a healthy cynicism to any quick fix. Like the underground circulation of cartoons in offices, school humor shows the intrusion of reform efforts. See *Everly's Laws of the Effect of Empirical Information on Educational Curriculum Reform*: "Law I: If it makes too much sense, educators will be suspicious of it" (Everly, 1987). Every researcher who plans conducting a study to prove the superiority of computerized instruction should read that

References

Apple, M. W. (1988). Teachers and texts: A political economy of class and gender relations in education. New York: Routledge.

Apple, M. W. & Weis, L. (1983). Ideology and practice in schooling. Philadelphia, PA: Temple University Press.

Barthes. R. (1973). Mythologies. London: Paladin Grafton Books.

Bradley, G. (1989). Computers and the psychosocial work environment. London: Taylor & Francis.

Cherryholmes, C. H. (1988). Power and criticism: Poststructural investigations in education. New York: Teachers College Press.

Ellsworth, E. (1989). Educational media, ideology, and the presentation of knowledge through popular cultural forms. In H. A. Giroux & R. I. Simon (Eds.) Popular culture, schooling and everyday life (pp. 47-66). Granby, MA: Bergin & Garvey.

Everly, P. A. (1987). Everly's laws of the effect of empirical information on educational curriculum reform. English Journal, 76 (1), 68.

Gibson, R. (1986). Critical theory and teaching. London: Hodder and Stoughton.

Giroux, H. A. (1983). Theory and resistance in education: A pedagogy for the opposition. New York: Bergin & Garvey.

Giroux, H. A. (1988). Teachers as intellectuals: Toward a critical pedagogy of learning. Granby, MA: Bergin & Garvey.

Giroux, H. A. & Simon, R. I. (Eds.) (1989). Popular culture, schooling and everyday life. Granby, MA: Bergin & Garvey.

Kidder, T. (1989). Among schoolchildren. New York: Avon.

Norman, D. A. (1988). The psychology of everyday things. New York: Basic Books.

Scalet, E. (1987). VDT health and safety: Issues and solutions. Lawrence KS: Ergosyst Associates.

So, you think you know "Technology"? (1988). Technical Communication 35 (1), 30.

Stellman, J. & Henifin, M. S. (1989). Office work can be dangerous to your health: A handbook of safety hazards and what you can do about them. New York: Fawcett Crest.

Taylor, W. D. & Johnsen, J. B. (1986). Resisting technological momentum. In J. A. Culbertson & L. L. Cunningham (Eds.) Microcomputers in education: Eighty-fifth year book of the National Society for the Study of Education (pp. 216-233). Chicago, IL: University of Chicago Press.

Tucker, M. (1985). Computers in the schools: What revolution? Journal of Communication, 35 (4), 12-23.

Zuboff, S. (1988). In the age of the smart machine: The future of work and power. New York: Basic Books.

short, informal lesson on the sociology of educational innovations. A current bestseller, *Among Schoolchildren*, describes nine months of observation in a fifth-grade class by an author previously recognized for writing about computers (Kidder, 1989). Computers are absent from this book, suggesting computers may be overrated and their presence in schools exaggerated.

- Norman, a cognitive scientist and author of *The Psychology of Everyday Things*, believes that badly designed hardware and software teach people to be helpless when working with technological items (1988, p. 40-43). Sometimes technical things seem to have been made for another species (p. 187). Given these sort of social limitations, schools prepare the workers of tomorrow for using computers and for learning from computerized instruction on the job. In human factors engineering this is called fitting the people to the machine. It is the last resort for overcoming inadequate design. The cognitive science viewpoint and the ergonomic viewpoint converge in that things need fixing, not people.
- Computers in schools are seen as TECHNOLOGY (Tucker, 1985). Therefore they are more desirable than other educational resources such as replacement textbooks (Taylor & Johnsen, 1986).

Synthesis

The abstract discussion of computers in education by scholars delivering conference papers may conflict ideologically with the practice of using computers in education. Giroux's insights into schooling in *Theory and Resistance in Education* can be adapted to provide an explanation (1983, p. 46):

1. Computers in education cannot be analyzed apart from the socioeconomic context in which they are situated.
2. Computers in education are politically involved in the construction and control of discourse, meaning and subjectivities.
3. The commonsense values and beliefs that guide and structure computers in education are not *a priori* universals but social constructions based on specific normative and political assumptions.

In synthesis, the data reported here give depth to the picture of computers in society, in work and in schools. The prices have dropped but computer corporations sell to schools, as they do to any other customer, to increase profits for themselves. Computerizing is a vehicle for social stratification. Computers are not easy to use and are hard to learn to use well. Schools play an important role in transmitting knowledge about how to use computers and by that increase sales and the rate of adoption. Workers and teachers object to the innovation in part because of design problems: computers do not always work well, are not always useful and can be a hindrance to getting things done. Some schools may be squandering money on computers instead of spending it on needed supplies. Where the quality of education was once rated on the ratio of teachers to students, it is now rated on the ratio of computers to students.

Title:

Survey of Instructional Design Needs and Competencies

Authors:

Barbara Seels
Z. Glasgow

Academic programs in instructional technology must serve multiple settings, such as business and industry, schools, government and the health professions. In some of these settings graduate degrees in instructional design (ID) are not perceived as a stepping stone to advancement. Thus, while academic programs in instructional design have a growing body of potential students, they may not be perceived as able to meet the needs of these students. One major problem of academic programs is determining what needs can be met and justifying the inclusion of these needs in the curriculum.

The faculty of academic programs have to be able to characterize the regional job market for ID professionals and the skills graduates must have to enter or advance in that market. To solve this problem the authors developed a questionaire survey that differs in many ways from the usual needs assessment in this area. The survey combines needs assessment and analysis of instructional requirements. Concurrently gathering information on jobs and task requirements proved to be a successful method for relating positions in the field to academic programs.

The questionaire was administered in the Western Pennsylvania region which was defined as covering from Morgantown, West Virginia in the south to Erie, Pennsylvania in the north and from Wheeling, West Virginia in the west to Johnstown and Indiana, Pennsylvania in the east. The questionaire gathered data on demographics of ID positions, on the background and competencies required for these positions and on the career objectives of ID professionals. There were also two questions about how a university or college could best respond to their ID needs. Questions 1-8 collected demographic and job information. Questions 9-17 covered competencies and task requirements. Questions 18-20 addressed career objectives and opportunities for professional development.

The task requirements section (questions 9-17) was based on a method developed by Fredrickson, Hawley, Whitemore and Wood (1980) for the Department of the Air Force. This method of task analysis identifies seven task criticality dimensions: 1. Learning Difficulty, 2. Performance Difficulty, 3. Consequence of Inadequate Performance, 4. Immediacy of Performance, 5. Availability of Assistance, 6. Frequency of Performance, 7. Proficiency Decay Rate. Each dimension is accompanied by a three point scale, such as easy-moderate-hard, low-moderate-high or infrequently-occasionally-frequently. The points on the scale are further defined by phrases, for example, easy-has little or no effect, moderate-could degrade or delay, hard-would result in failure. A numerical value is given to a point on the scale.

These numerical values are used to calculate recommended training requirements including certification or qualification (formal training), on-the-job training (OJT), proficiency maintenance, reduced time for training or elimination from training. For example, if learning difficulty and consequence of performance are rated low and availability of assistance is rated high, it is recommended that these tasks be eliminated from training because they will not be important to success.

Initially, the questionaire included each of the criticality dimensions followed by a list of 16 competencies. Seven dimensions with 16 items each to respond to on a five point scale made the questionaire too long to be feasible. Even though respondents were directed to indicate the level of criticality for only those competencies they checked as relevant to their job, the procedure was too time-consuming. As a result, the second draft of the questionaire was based on the four criticality dimensions deemed by the authors to be most relevant to the study. These four criticality dimensions were restated as:

1. Rate the difficulty of this task.
2. Were you expected to perform this task immediately upon starting your present position?
3. Are there people in your organization to whom you can go for help with this task?
4. What happens if you perform this task improperly?

The revised questionaire simplified the answers required. Answer formats varied from checking one of three or four phrases to checking yes or no. The revised questionaire also allowed for an open-ended response to ways the university can best meet your needs. Because there were changes in the dimensions, the scales and the mode of responding, it was no longer possible to use the questionaire to numerically determine the recommended type of training.

Questions 9-17 were also based on an analysis of ID competencies. The NSPI/AECT (National Society for Performance and Instruction/Association for Educational Communications and Technology) list of ID competencies was a starting point. Sixteen items on this list were given to beginning students in ID who were asked to write the meaning of each competency. This procedure allowed identification of words and phrases confusing to those who might be in the field doing the tasks but who were unfamiliar with some of the terminology of the field. One competency that proved to be very confusing as stated was "analyze the structural characteristics of jobs, tasks, and content." The list was then reduced to the most essential items in the opinion of the authors and confusing items were restated. The final list of nine competencies for the questionaire was:

a. establishing educational goals
b. doing task analysis
c. determining learner characteristics
d. writing objectives
e. selecting strategies for instruction
f. developing media
g. evaluating instruction
h. managing ID projects
i. promoting adoption of instructional programs

A mailing list for the survey was developed from the lists of four professional associations with members in the region. These associations were: Pennsylvania Association for Educational Communications and Technology, National Society for Performance and Instruction, American Society for Training and Development and Pennsylvania Association for

Supervision and Curriculum Development. All regional members of the first two associations were included. The list from the Pittsburgh Chapter of the American Society for Training and Development was culled by selecting only those members whose job titles or organizations indicated relationship to ID tasks. For example, all members from Development Dimensions International and Applied Science Associates were included. Members with titles such as Training Development Specialist or Manager were included; while those with titles such as Human Resource Specialist or Manager were not. If the authors knew a member was involved with ID, that person was also included. Regional members of the Pennsylvania Association for Supervision and Curriculum were so numerous that list was sampled. Of the 104 local members of that association 63 were sent questionaires. To reduce the list all the names from one district were put on cards and one card was drawn randomly. A district with only one member received a questionaire; a district with many members also received only one questionaire.

The cover letter began by defining ID as a "step-by-step systematic process for developing teaching-learning situations. This process consists of procedures and guidelines that help the designer determine learning objectives, plan instructional events, and measure learning outcomes." Those receiving the questionaire were then asked to complete it if they were involved with ID tasks or positions. They were told this information would be used to plan professional development opportunities for the region. Those who decided they did not supervise or perform ID tasks were asked to pass the questionaire to someone who did, if possible.

Two hundred and sixty questionaires were mailed. A post-card reminder followed three weeks later. Sixty-three questionaires were returned making a return rate of about 25%. Of the surveys returned 47.6% or 30 were from school settings, 42.8% or 27 were from other settings, such as business and industry, government, health, consulting. Six or 9.5% were from higher education. It was decided that the return rate from higher education was too small to provide useable data. Because the results apply to the Western Pennsylvania region, they are not generalizable, but some may be of interest. The data was analyzed for the whole population and then broken down by settings: schools, other and higher education. Seven findings of interest are:

1. One third of those employed in ID tasks have no one to help them with these tasks.

2. The task of developing media is the least critical task and is the task done the least.

3. Generally, "doing the task improperly" had a moderate to serious effect. Whether the effect was moderate or serious depended on the setting and competency.

4. Therefore, the most important tasks depend on the setting. For school personnel they are evaluating, managing and diffusing. For other settings (industry, health, etc.) the most important task is establishing goals.

5. There is more perceived instructional design being accomplished in schools than is generally assumed

6. Much instructional design is done by people not prepared in the instructional design and technology field. (The questionaire did not measure quality of performance.

7. About a third of those responding did not perceive task analysis to be critical.

A more complete report on the results is available from the authors. The questionaire is given at the end of this article. So far the validity of the questionaire has been content validity based on the process of development and review by two ID experts and a measurement and evaluation expert. The reliability has not been established. It has been used only in this study. Those who use the questionaire are requested to credit the authors and to share their results with the authors. This will help establish validity and reliability beyond the Western Pennsylvania region.

References

Bratton, B. (1984). Professional certification: Will it become a reality? *Performance and Instruction*, 23, 4-7.

Fredrickson, E.W., Hawley, J., Whitemore, P., and Wood, M.M. (1980). *Air defense training development decision aid (TDDA): Model extension and research requirement, final report.* Contract No. MDA 903-80-C-0160. Army Research Institute, Defense Supply Institute. Butler, PA: Applied Science Associates.

Seels, B. and Glasgow, Z. (1990). *Exercises in Instructional Design.* Columbus: Merrill Publishing Co.

Seels, B. and Glasgow, Z. (1990). *Instructor's Manual for Exercises in Instructional Design.* Columbus: Merrill Publishing Co.

Appendix: The Questionaire

SURVEY ON COMPETENCY REQUIREMENTS
FOR INSTRUCTIONAL DESIGN POSITIONS

Directions: For each item below, please check the appropriate response.

1. In what setting do you work?

    ______ School System
    ______ Business and Industry
    ______ Higher Education
    ______ Health Agency
    ______ Government
    ______ Consulting Firm
    ______ Other, specify

2. How many employees in your organization spend at least 50% of the time on ID tasks? ____________

3. During the next five years do you expect the number of ID positions in your organization to:

    ______ Increase ______ Stay the same ______ Decrease

4. How many years have you been employed in ID?

    ______ 2 years or less
    ______ 3 to 5 years
    ______ 6 to 10 years
    ______ 11 to 20 years
    ______ more than 20 years

5. How would you describe your position within your organization?

    ______ Staff ______ Supervisor ______ Manager

6. What portion of your time is devoted to performing or managing ID?

    ______ 20% or less
    ______ 21 to 40%
    ______ 41 to 60%
    ______ 61 to 80%
    ______ 81 to 100%

7. Check your level of education and indicate your major area of study at that level.

    ______ Associate/Technical Degree ______________________________
    ______ Bachelor's Degree ______________________________
    ______ Master's Degree ______________________________
    ______ Doctoral Degree ______________________________

8. Rate the relevance of each of your educational degrees to your present position.

	None	Some	Great deal
Associate/technical Degree	1	2	3
Bachelor's Degree	1	2	3
Master's Degree	1	2	3
Doctorate	1	2	3

Directions: Questions 9 - 17 present a variety of tasks people are asked to do in instructional design positions. Each question asks first if you perform the task. If you answer yes, you are given a set of questions on the task.

9. Are you presently involved in establishing instructional goals?

______ Yes ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of this task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes ______ No

c. Are there people in your organization to whom you can go for help with this task?

______ Yes ______ No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

10. Are you presently involved in doing task analysis?

______ Yes ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do so
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present task?

______ Yes ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

11. Are you presently involved in determining learner characteristics?

______ Yes ______No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes   ______ No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

12. Are you presently involved in writing objectives?

______ Yes   ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes   ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes   ______ No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

13. Are you presently involved in selecting strategies for conducting instruction?

______ Yes   ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes   ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes   ______ No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

14. Are you presently involved in developing media?

______ Yes   ______ No

If you answered yes to this question, please complete items a through e.

a. What kind of media do you develop? ____________________________________

____

b. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

c. Were you expected to perform this task immediately upon starting your present position?

______ Yes ______ No

d. Are there people in your organization to whom you can go for help to do this task?

______ Yes ______ No

e. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

15. Are you presently involved in evaluating instruction?

______ Yes ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes ______ No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

16. Are you presently involved in managing ID projects?

______ Yes ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes ______ No

d. What happens if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

17. Are you presently involved in promoting adoption of instructional programs?

______ Yes ______ No

If you answered yes to this question, please complete items a through d.

a. Rate the difficulty of the task.

______ Easy to do
______ Moderately difficult to do
______ Difficult to do

b. Were you expected to perform this task immediately upon starting your present position?

______ Yes ______ No

c. Are there people in your organization to whom you can go for help to do this task?

______ Yes ______ No

d. What happen if you perform this task improperly?

______ has no effect
______ has minimal effect
______ has moderate effect
______ has serious effect

18. Which of the following best describes your career objectives for the next 3 years?

______ Maintain skills in current position
______ Upgrade ID skills for current position
______ Add new skills to obtain new ID position
______ Leave ID field

19. Check any of the following which you feel would be useful for achieving your career objectives.

______ Certification courses
______ Short workshops (1/2 day)
______ Longer workshops (1-3days)
______ Degree programs
______ Internships
______ Other, specify ______________________________

20. What kind of professional opportunities in instructional design would you like to see offered at the university level?

RETURN TO:

Dr. Barbara Seels
4A16 Forbes Quadrangle
University of Pittsburgh
Pittsburgh, PA 15260

1991

AUTHOR INDEX

Authors	Descriptors	YR PG
Anglin, G	instructional design	91/1
Anglin, G	networks	91/1
Arenz, B	cooperative learning	91/21
Arenz, B	problem solving	91/21
Arnone, M	interactive video	91/45
Arnone, M	learner control	91/45
Aust, R	navigation tools	91/69
Aust, R	networks	91/69
Berry, E	hemispheric laterality	91/929
Berry, E	realism	91/929
Berry, L	cognitive style	91/289
Berry, L	memory	91/83,91
Berry, L	realism	91/83
Berry, L	visuals	91/91
Berry, L	visual recall	91/289
Bloom, E	expert systems	91/593
Bloom, E	tutoring	91/593
Bork, A	education	91/349
Bork, A	technology	91/349
Borsook, T	interactivity	91/103
Borsook, T	instruction	91/103
Caffarella, E	instructional design	91/119
Caffarella, E	instructional technology	91/119
Cardinale, L	information processing	91/131
Cardinale, L	learning	91/131
Christensen, D	instructional development	91/885
Christensen, D	expert systems	91/885
Cleveland, E	telecommunications	91/723
Cleveland, E	tutoring	91/723
Cole, P	cognitive apprenticeship	91/945
Cole, P	instructional design	91/945
Cook, P	authoring	91/941
Cook, P	expert systems	91/941
Coorough, R	anxiety	91/147
Coorough, R	computer-assisted instruction	91/147
Dempsey, J	evaluation	91/167
Dempsey, J	qualitative research	91/167
De Vaney, A	gender	91/191
De Vaney, A	television	91/191
Elenes, A	gender	91/191
Elenes, A	television	91/191
Frick, T	expert systems	91/213
Frick, T	response theory	91/213
George, R	qualitative research	91/235
George, R	research methodology	91/235
Gillis, P	expert systems	91/593
Gillis, P	tutoring	91/593
Glasgow, Z	instructional design	91/995
Glasgow, Z	needs assessment	91/995
Goodrum, D	design	91/245
Goodrum, D	process tools	91/245
Grabowski, B	learner control	91/45
Grabowski, B	interactive video	91/45
Green, E	authoring	91/941

Authors	Descriptors	YR PG
Green, E	expert systems	91/941
Harvey, F	hypermedia	91/843
Harvey, F	research	91/843
Higginbotham-Wheat, N	computer-based instruction	91/261
Higginbotham-Wheat, N	systems	91/261
Igoe, A	academic achievement	91/275
Igoe, A	motivation	91/275
Januszewski, A	educational technology	91/395
Januszewski, A	theory	91/395
Jesky, R	cognitive style	91/289
Jesky, R	visual recall	91/289
Jones, M	cognitive theory	91/625
Jones, M	instructional strategies	91/625
King, J	communication	91/297
King, J	educational technology	91/297
Klayder, J	navigation tools	91/69
Klayder, J	networks	91/69
Klein, J	learner control	91/481
Klein, J	simulation	91/481
Knupfer, N	design	91/313,383
Knupfer, N	desktop publishing	91/313,383
Knupfer, N	educational computing	91/331
Knupfer, N	pedagogy	91/331
Knuth, R	design	91/245
Knuth, R	process tools	91/245
Koetting, J	educational technology	91/395
Koetting, J	theory	91/395
Koetting, R	curriculum theory	91/381
Koetting, R	educational technology	91/381
Lee, D	computer- assisted Instruction	91/411
Lee, D	feedback	91/411
Lee, M	loci of control	91/433
Lee, M	metacognition	91/433
Legree, P	expert systems	91/593
Legree, P	tutoring	91/593
Lowry, W	achievement	91/461
Lowry, W	videodisc	91/461
Lu, M	educational media	91/521
Lu, M	utilization	91/521
Mattoon, J	learner control	91/481
Mattoon, J	simulation	91/481
Mayton, G	animation	91/549
Mayton, G	computer-based instruction	91/549
McIsaac, M	design	91/313,383
McIsaac, M	desktop publishing	91/313,383
McIsaac, M	distance education	91/343
McIsaac, M	networks	91/343
Misanchuk, E	audit	91/499
Misanchuk, E	interactive media	91/499
Morlan, J	educational media	91/521
Morlan, J	utilization	91/521
Morrison, G	tutoring	91/723
Morrison, G	telecommunications	91/723
Morrissey, K	learning	91/537

Authors	Descriptors	YR PG
Morrissey, K	interactive video	91/537
Muffoletto, R	empowerment	91/341
Muffoletto, R	computers	91/341
Nelson, W	cognitive theory	91/611
Nelson, W	tutoring	91/611
Newby, T	concept development	91/561
Newby, T	learning	91/561
Norris, C	education	91/581
Norris, C	restructuring	91/581
Orey, M	cognitive theory	91/593,625
Orey, M	expert systems	91/593
Orey, M	instructional strategies	91/625
Orey, M	tutoring	91/593,611
Okey, J	cognitive theory	91/625
Okey, J	instructional strategies	91/625
Reigeluth, C	education	91/581
Reigeluth, C	restructuring	91/581
Richey, R	attitude	91/635
Richey, R	training	91/635
Rieber, L	animation	91/681
Rieber, L	constructionism	91/691
Rieber, L	microworlds	91/691
Rieber, L	visuals	91/681
Rodriguez, S	instructional strategies	91/709
Rodriguez, S	learning	91/709
Ross, S	computers	91/741
Ross, S	telecommunications	91/723
Ross, S	tutoring	91/723
Ross, S	writing	91/741
Russell, F	distance education	91/759
Russell, F	satellite-delivered instruction	91/759
Sachs, S	instructional design	91/119
Sachs, S	instructional technology	91/119
Savenye, W	computer- assisted instruction	91/411
Savenye, W	feedback	91/411
Schlenk, G	instructional development	91/779
Schlenk, G	research	91/779
Schwier, R	audit	91/499
Schwier, R	interactive media	91/499
Seels, B	instructional design	91/995
Seels, B	needs assessment	91/995
Shih, M	distance education	91/795
Shih, M	models	91/795
Shrock, S	evaluation	91/817
Shrock, S	instructional design	91/817
Shrock, S	instructional development	91/779
Shrock, S	research	91/779
Smith, L	computers	91/741
Smith, L	telecommunications	91/723
Smith, L	tutoring	91/723
Smith, L	writing	91/741
Smith, M	design	91/803
Smith, M	visuals	91/803
Smith, P	computer- assisted instruction	91/411

Authors	Descriptors	YR PG
Smith, P	design	91/803
Smith, P	feedback	91/411
Smith, P	visuals	91/803
Stanley, L	cognitive theory	91/625
Stanley, L	instructional strategies	91/625
Stepich, D	concept development	91/561
Stepich, D	learning	91/561
Stepp, S	evaluation	91/817
Stepp, S	instructional design	91/817
Story, W	hypermedia	91/843
Story, W	research	91/843
Streibel, M	instructional design	91/849
Streibel, M	theory	91/849
Tennyson, R	expert systems	91/885
Tennyson, R	instructional development	91/885
Thorkildsen, R	achievement	91/461
Thorkildsen, R	videodisc	91/461
Thurman, R	learner control	91/481
Thurman, R	simulation	91/481
Towers, R	instructional design	91/1
Towers, R	networks	91/1
Tripp, S	instructional design	91/917
Tripp, S	theory	91/917
Tucker, S	evaluation	91/167
Tucker, S	qualitative research	91/167
Waltz, E	hemispheric laterality	91/929
Waltz, E	realism	91/929
Wilkins, D	authoring	91/941
Wilkins, D	expert systems	91/941
Wilson, B	cognitive apprenticeship	91/945
Wilson, B	instructional design	91/945
Wolcott, L	distance education	91/957
Wolcott, L	teaching	91/957
Woodson, E	computers	91/741
Woodson, E	writing	91/741
Yeaman, A	computers	91/989
Yeaman, A	education	91/989
Yeaman, A	educational technology	91/979
Yeaman, A	theory	91/979
Zvacek, S	distance education	91/795
Zvacek, S	models	91/795

1991

DESCRIPTOR INDEX

Descriptors	Authors	YR PG
academic achievement	Igoe, A	91/275
achievement	Lowry, W	91/461
achievement	Thorkildsen, R	91/461
animation	Mayton, G	91/549
animation	Rieber, L	91/681
anxiety	Coorough, R	91/147
attitude	Richey, R	91/635
audit	Misanchuk, E	91/499
audit	Schwier, R	91/499
authoring	Cook, P	91/941
authoring	Green, E	91/941
authoring	Wilkins, D	91/941
cognitive apprenticeship	Cole, P	91/945
cognitive apprenticeship	Wilson, B	91/945
cognitive style	Berry, L	91/289
cognitive style	Jesky, R	91/289
cognitive theory	Jones, M	91/625
cognitive theory	Nelson, W	91/611
cognitive theory	Okey, J	91/625
cognitive theory	Orey, M	91/593,625
cognitive theory	Stanley, L	91/625
communication	King, J	91/297
computers	Muffoletto, R	91/341
computers	Ross, S	91/741
computers	Smith, L	91/741
computers	Woodson, E	91/741
computers	Yeaman, A	91/989
computer-assisted instruction	Coorough, R	91/147
computer- assisted instruction	Lee, D	91/411
computer- assisted instruction	Savenye, W	91/411
computer- assisted instruction	Smith, P	91/411
computer-based instruction	Higginbotham-Wheat, N	91/261
computer-based instruction	Mayton, G	91/549
concept development	Newby, T	91/561
concept development	Stepich, D	91/561
constructionism	Rieber, L	91/691
cooperative learning	Arenz, B	91/21
curriculum theory	Koetting, R	91/381
desktop publishing	Knupfer, N	91/313,383
desktop publishing	McIsaac, M	91/313,383
design	Goodrum, D	91/245
design	Knupfer, N	91/313,383
design	Knuth, R	91/245
design	McIsaac, M	91/313,383
design	Smith, M	91/803
design	Smith, P	91/803
distance education	McIsaac, M	91/343
distance education	Russell, F	91/759
distance education	Shih, M	91/795
distance education	Shih, M	91/795
distance education	Wolcott, L	91/957
distance education	Zvacek, S	91/795
education	Bork, A	91/349
education	Norris, C	91/581

Descriptors	Authors	YR PG
education	Reigeluth, C	91/581
education	Yeaman, A	91/989
educational computing	Knupfer, N	91/331
educational media	Morlan, J	91/521
educational media	Lu, M	91/521
educational technology	Januszewski, A	91/395
educational technology	Koetting, J	91/395
educational technology	Koetting, R	91/381
educational technology	King, J	91/297
educational technology	Yeaman, A	91/979
empowerment	Muffoletto, R	91/341
evaluation	Dempsey, J	91/167
evaluation	Shrock, S	91/817
evaluation	Stepp, S	91/817
evaluation	Tucker, S	91/167
expert systems	Bloom, E	91/593
expert systems	Christensen, D	91/885
expert systems	Cook, P	91/941
expert systems	Frick, T	91/213
expert systems	Green, E	91/941
expert systems	Gillis, P	91/593
expert systems	Legree, P	91/593
expert systems	Orey, M	91/593
expert systems	Tennyson, R	91/885
expert systems	Wilkins, D	91/941
feedback	Lee, D	91/411
feedback	Savenye, W	91/411
feedback	Smith, P	91/411
gender	De Vaney, A	91/191
gender	Elenes, A	91/191
hemispheric laterality	Berry, E	91/929
hemispheric laterality	Waltz, E	91/929
hypermedia	Harvey, F	91/843
hypermedia	Story, W	91/843
information processing	Cardinale, L	91/131
instruction	Borsook, T	91/103
instructional design	Anglin, G	91/1
nstructional design	Caffarella, E	91/119
instructional design	Cole, P	91/945
instructional design	Glasgow, Z	91/995
instructional design	Sachs, S	91/119
instructional design	Seels, B	91/995
instructional design	Shrock, S	91/817
instructional design	Stepp, S	91/817
instructional design	Streibel, M	91/849
instructional design	Towers, R	91/1
instructional design	Tripp, S	91/917
instructional design	Wilson, B	91/945
instructional development	Christensen, D	91/885
instructional development	Schlenk, G	91/779
instructional development	Shrock, S	91/779
instructional development	Tennyson, R	91/885
instructional strategies	Jones, M	91/625
instructional strategies	Okey, J	91/625

Descriptors	Authors	YR PG
instructional strategies	Orey, M	91/625
instructional strategies	Rodriguez, S	91/709
instructional strategies	Stanley, L	91/625
instructional technology	Caffarella, E	91/119
instructional technology	Sachs, S	91/119
interactivity	Borsook, T	91/103
interactive media	Misanchuk, E	91/499
interactive media	Schwier, R	91/499
interactive video	Arnone, M	91/45
interactive video	Grabowski, B	91/45
interactive video	Morrissey, K	91/537
learner control	Arnone, M	91/45
learner control	Grabowski, B	91/45
learner control	Klein, J	91/481
learner control	Mattoon, J	91/481
learner control	Thurman, R	91/481
learning	Cardinale, L	91/131
learning	Morrissey, K	91/527
learning	Newby, T	91/561
learning	Rodriguez, S	91/709
learning	Stepich, D	91/561
loci of control	Lee, M	91/433
memory	Berry, L	91/83,91
metacognition	Lee, M	91/433
microworlds	Rieber, L	91/691
models	Shih, M	91/795
models	Zvacek, S	91/795
motivation	Igoe, A	91/275
navigation tools	Aust, R	91/69
navigation tools	Klayder, J	91/69
needs assessment	Glasgow, Z	91/995
needs assessment	Seels, B	91/995
networks	Anglin, G	91/1
networks	Aust, R	91/69
networks	Klayder, J	91/69
networks	McIsaac, M	91/343
networks	Towers, R	91/1
pedagogy	Knupfer, N	91/331
problem solving	Arenz, B	91/21
process tools	Goodrum, D	91/245
process tools	Knuth, R	91/245
qualitative research	Dempsey, J	91/167
qualitative research	George, R	91/235
qualitative research	Tucker, S	91/167
realism	Berry, E	91/929
realism	Berry, L	91/83
realism	Waltz, E	91/929
research	Harvey, F	91/843
research	Schlenk, G	91/779
research	Shrock, S	91/779
research	Story, W	91/843
research methodology	George, R	91/235
response theory	Frick, T	91/941
restructuring	Norris, C	91/581

Descriptors	Authors	YR PG
restructuring	Reigeluth, C	91/581
satellite-delivered instruction	Russell, F	91/759
simulation	Klein, J	91/481
simulation	Mattoon, J	91/481
simulation	Thurman, R	91/481
systems	Higginbotham-Wheat, N	91/261
teaching	Wolcott, L	91/957
technology	Bork, A	91/349
telecommunications	Cleveland, E	91/723
telecommunications	Morrison, G	91/723
telecommunications	Ross, S	91/723
telecommunications	Smith, L	91/723
television	De Vaney, A	91/191
television	Elenes, A	91/191
theory	Januszewski, A	91/395
theory	Koetting, J	91/395
theory	Streibel, M	91/849
theory	Tripp, S	91/917
theory	Yeaman, A	91/979
training	Richey, R	91/635
tutoring	Bloom, E	91/593
tutoring	Cleveland, E	91/723
tutoring	Gillis, P	91/593
tutoring	Legree, P	91/593
tutoring	Morrison, G	91/723
tutoring	Nelson, W	91/611
tutoring	Orey, M	91/593,611
tutoring	Ross, S	91/723
tutoring	Smith, L	91/723
utilization	Lu, M	91/521
utilization	Morlan, J	91/521
videodisc	Lowry, W	91/461
videodisc	Thorkildsen, R	91/461
visuals	Berry, L	91/91
visuals	Rieber, L	91/681
visuals	Smith, M	91/803
visuals	Smith, P	91/803
visual recall	Berry, L	91/289
visual recall	Jesky, R	91/289
writing	Ross, S	91/741
writing	Smith, L	91/741
writing	Woodson, E	91/741